SUPPLEMENTAL MATERIALS TO

"Mode-based Classifier: A Robust and Flexible Discriminant Analysis for High-dimensional Data"

University of International Business and Economics, Humboldt-Universitat zu Berlin, Brunel University and Renmin University of China

Supplementary Material

In this document, we first provide theoretical properties and justifications for Theorems 1-3, then we investigate the rate of convergence of the unimodal classifiers when p is fixed and $p \to \infty$. In section 2, we present the simulation settings for Figure 1 and Table 1. In section 3, we provide an illustrative example to demonstrate how to select σ_{11} in the mode-based classifiers. The additional simulation results for Examples 1-3 are listed in Section 4. Finally, we display some tables and figures for real data analysis.

S1 Proofs for Theorems 1-3

Lemma S1. *When $p = 1$, and assume the bandwidths are equal for two populations and take a Gaussian kernel, the probability of correct classification for the mode classifier takes the following simple form:*

if $\delta_1 \leq \delta_2$,

$$\Gamma_\sigma(\delta) = \pi_1 \mathcal{F}_1\big((\delta_1 + \delta_2)/2\big) + \pi_2\big\{1 - \mathcal{F}_2((\delta_1 + \delta_2)/2)\big\},$$

if $\delta_1 > \delta_2$,

$$\Gamma_\sigma(\delta) = \pi_1\big\{1 - \mathcal{F}_1((\delta_1+\delta_2)/2)\big\} + \pi_2\mathcal{F}_2\big((\delta_1+\delta_2)/2\big).$$

where δ_1 *and* δ_2 *are the true modes of the two populations.*

Proof. In the univariate case, since $\sigma_1 = \sigma_2$, $\Psi_1(z,\delta,\sigma) = \varphi(\frac{z-\delta_1}{\sigma})$ and $\Psi_2(z,\delta,\sigma) = \varphi(\frac{z-\delta_2}{\sigma})$. For the true modes δ_1 and δ_2, the integral (5) can be computed by splitting it into four parts according to the possible disjoint regions of the domain of Z with respect to δ_1 and δ_2, that is : (a) $a \leq \min\{\delta_1,\delta_2\}$, (b) $\delta_1 < z \leq \delta_2$, (c) $\delta_2 < z \leq \delta_1$ and (d) $z > \max\{\delta_1,\delta_2\}$.

For case (a), if $\delta_1 < \delta_2$, the integral becomes

$$\begin{aligned}\Gamma_a(\delta) &= \pi_1\int_{-\infty}^{\delta_1} I\Big\{\varphi\Big(\frac{z-\delta_1}{\sigma}\Big) > \varphi\Big(\frac{z-\delta_2}{\sigma}\Big)\Big\} f_1(z)dz = \pi_1\int_{-\infty}^{\delta_1} f_1(z)dz \\ &= \pi_1\mathcal{F}_1(\delta_1),\end{aligned}$$

otherwise, for $\delta_1 \geq \delta_2$, $\Gamma_a(\delta) = \pi_2\mathcal{F}_2(\delta_2)$.

In case (b), the integral is

$$\begin{aligned}\Gamma_b(\delta) &= \pi_1\int_{\delta_1}^{\delta_2} I\Big\{\varphi\Big(\frac{z-\delta_1}{\sigma}\Big) > \varphi\Big(\frac{z-\delta_2}{\sigma}\Big)\Big\} f_1(z)dz \\ &\quad +\pi_2\int_{\delta_1}^{\delta_2} I\Big\{\varphi\Big(\frac{z-\delta_1}{\sigma}\Big) \leq \varphi\Big(\frac{z-\delta_2}{\sigma}\Big)\Big\} f_2(z)dz \\ &= \pi_1\int_{\delta_1}^{(\delta_1+\delta_2)/2} f_1(z)dz + \pi_2\int_{(\delta_1+\delta_2)/2}^{\delta_2} f_2(z)dz \\ &= \pi_1\{\mathcal{F}_1((\delta_1+\delta_2)/2) - \mathcal{F}_1(\delta_1)\} + \pi_2\{\mathcal{F}_2(\delta_2) - \mathcal{F}_2((\delta_1+\delta_2)/2)\}.\end{aligned}$$

Similarly, for cases (c) and (d),

$$\Gamma_c(\delta) = \pi_1\{\mathcal{F}_1(\delta_1) - \mathcal{F}_1((\delta_1+\delta_2)/2)\} + \pi_2\{\mathcal{F}_2((\delta_1+\delta_2)/2) - \mathcal{F}_2(\delta_2)\}.$$

if $\delta_1 < \delta_2$, $\Gamma_d(\delta) = \pi_2(1-\mathcal{F}_2(\delta_2))$, and if $\delta_1 \geq \delta_2$, $\Gamma_d(\delta) = \pi_1(1-\mathcal{F}_1(\delta_1))$. When $\delta_1 \leq \delta_2$, $\Gamma_\sigma(\delta)$ is the sum of $\Gamma_a(\delta)$, $\Gamma_b(\delta)$ and $\Gamma_d(\delta)$ corresponding to disjoint domain regions of Z:

$$\Gamma_\sigma(\delta) = \pi_1\mathcal{F}_1\big((\delta_1+\delta_2)/2\big) + \pi_2\big\{1-\mathcal{F}_2((\delta_1+\delta_2)/2\big\},$$

and when $\delta_1 > \delta_2$,

$$\Gamma_\sigma(\delta) = \pi_1\big\{1-\mathcal{F}_1((\delta_1+\delta_2)/2)\big\} + \pi_2\mathcal{F}_2\big((\delta_1+\delta_2)/2\big).$$

This completes the proof. □

To prove Theorems 1-3, we need the following lemmas. The proofs are based on the corrected modal distance, for the simplified one, all the lemmas and theorems hold by replacing σ_{rj} by 1, for $r = 1, 2$ and $j = 1, \dots, p$.

Lemma S2. *Under the assumption A1 and A2, for $j \in \{1, \dots, p\}$, the following equation holds for $\delta_{1j}, \delta_{2j} \in \mathbb{R}$ and $\sigma_{1j}, \sigma_{2j} \in (0, \infty)$,*

$$|\Psi_{1j}(z_j,\sigma_{1j}) - \Psi_{2j}(z_j,\sigma_{2j})| \leq \frac{c_1}{\sigma_{1j}}|z_j|\left|\frac{1}{\sigma_{1j}} - \frac{1}{\sigma_{2j}}\right| + \frac{c_1}{\sigma_{1j}}\left|\frac{\delta_{1j}}{\sigma_{1j}} - \frac{\delta_{2j}}{\sigma_{1j}}\right| + c_0\left|\frac{1}{\sigma_{1j}} - \frac{1}{\sigma_{2j}}\right|.$$

Proof. Proof of Lemma S2. By definition,

$$
\begin{aligned}
&|\Psi_{1j}(z_j,\sigma_{1j})-\Psi_{2j}(z_j,\sigma_{2j})| = \left|\frac{1}{\sigma_{1j}}K\left(\frac{z_j-\delta_{1j}(\sigma_{1j})}{\sigma_{1j}}\right)-\frac{1}{\sigma_{2j}}K\left(\frac{z_j-\delta_{2j}(\sigma_{2j})}{\sigma_{2j}}\right)\right| \\
= \ & \frac{1}{\sigma_{1j}\sigma_{2j}}\left|\sigma_{2j}K\left(\frac{z_j-\delta_{1j}(\sigma_{1j})}{\sigma_{1j}}\right)-\sigma_{2j}K\left(\frac{z_j-\delta_{2j}(\sigma_{2j})}{\sigma_{2j}}\right)+\sigma_{2j}K\left(\frac{z_j-\delta_{2j}(\sigma_{2j})}{\sigma_{2j}}\right)\right. \\
& \left.-\sigma_{1j}K\left(\frac{z_j-\delta_{2j}(\sigma_{2j})}{\sigma_{2j}}\right)\right| \\
= \ & \frac{1}{\sigma_{1j}}\left|K\left(\frac{z_j-\delta_{1j}(\sigma_{1j})}{\sigma_{1j}}\right)-K\left(\frac{z_j-\delta_{2j}(\sigma_{2j})}{\sigma_{2j}}\right)\right|+\left|\frac{\sigma_{1j}-\sigma_{2j}}{\sigma_{1j}\sigma_{2j}}\right|K\left(\frac{z_j-\delta_{2j}(\sigma_{2j})}{\sigma_{2j}}\right) \\
\leq \ & \frac{1}{\sigma_{1j}}\left|K'\left(\frac{z_j-\delta_{2j}(\sigma_{2j})}{\sigma_{2j}}\right)\right|\left|\left(\frac{z_j-\delta_{1j}(\sigma_{1j})}{\sigma_{1j}}\right)-\left(\frac{z_j-\delta_{2j}(\sigma_{2j})}{\sigma_{2j}}\right)\right|+c_0\left|\frac{1}{\sigma_{1j}}-\frac{1}{\sigma_{2j}}\right| \\
\leq \ & \frac{c_1}{\sigma_{1j}}|z_j|\left|\frac{1}{\sigma_{1j}}-\frac{1}{\sigma_{2j}}\right|+\frac{c_1}{\sigma_{1j}}\left|\frac{\delta_{1j}}{\sigma_{1j}}-\frac{\delta_{2j}}{\sigma_{1j}}\right|+c_0\left|\frac{1}{\sigma_{1j}}-\frac{1}{\sigma_{2j}}\right|
\end{aligned}
$$

The first and second inequality holds due to the assumption that $|K(\cdot)| \leq c_0$ and $|K'(\cdot)| \leq c_1$, respectively. This completes the proof. □

Remark 1. It follows naturally that

$$|\Psi_{1j}(z_j,\sigma_{1j})-\Psi_{2j}(z_j,\theta\sigma_{1j})| \leq \frac{|\theta-1|}{\theta\sigma_{1j}}\left(c_0+\frac{c_1}{\sigma_{1j}}|z_j|\right)+\frac{c_1}{\theta\sigma_{1j}^2}|\theta\delta_{1j}-\delta_{2j}|. \tag{S1.1}$$

Remark 2. By the same strategy, we can obtain that

$$|\Psi_{1j}(z_j,\sigma_1^*)-\Psi_{1j}(z_j,\sigma_1)| \leq c_1|z_j-\delta_{1j}|\left|\frac{1}{\sigma_1^*}-\frac{1}{\sigma_1}\right|+c_0\left|\frac{1}{\sigma_1^*}-\frac{1}{\sigma_1}\right|. \tag{S1.2}$$

and

$$|\Psi_{2j}(z_j,\sigma_2^*)-\Psi_{2j}(z_j,\sigma_2)| \leq c_1|z_j-\delta_{2j}|\left|\frac{1}{\sigma_2^*}-\frac{1}{\sigma_2}\right|+c_0\left|\frac{1}{\sigma_2^*}-\frac{1}{\sigma_2}\right|. \tag{S1.3}$$

Lemma S3. *Under assumptions (A1) and (A2), for $j \in \{1, \ldots, p\}$, the following equation holds for $\delta_{2j} \in \mathbb{R}$ and $\sigma_{1j} \in (0, \infty)$,*

$$\sigma_{1j}|\Psi_{2j}(z_j, \theta_1\sigma_{1j}) - \Psi_{2j}(z_j, \theta_2\sigma_{1j})| \leq c_0\left|\frac{1}{\theta_1} - \frac{1}{\theta_2}\right| + \frac{c_1|z_j|}{\sigma_{1j}\theta_2}\left|\frac{1}{\theta_1} - \frac{1}{\theta_2}\right| + \frac{c_1}{\theta_2}\left|\frac{\delta_{2j}(\theta_1\sigma_{1j})}{\theta_1\sigma_{1j}} - \frac{\delta_{2j}(\theta_2\sigma_{1j})}{\theta_2\sigma_{1j}}\right|. \tag{S1.4}$$

Proof. Proof of Lemma S3 works analogy of proof of Lemma S2. □

Lemma S4. *Under assumptions (A1)-(A3), then $\hat{\delta}_{ij,\sigma_{ij,n}} \to \delta_{ij}$ a.s., and $\hat{f}_{ij,\sigma_{ij,n}}(\hat{\delta}_{ij,\sigma_{ij,n}}) \to f_{ij}(\delta_{ij})$ a.s.*

Proof. To simplify the proof, we abbreviate the notations δ_{ij} as δ, $\sigma_{ij,n}$ as σ_n, f_{ij} as f. For independent and identically distributed samples $U_1, \ldots, U_n$ i.i.d. with density f, let

$$\hat{f}_{n,\sigma_n} = \frac{1}{n\sigma_n}\sum_{i=1}^{n} K\left(\frac{t - U_i}{\sigma_n}\right).$$

First, we show the following. Let a_n and b_n be fixed sequences of numbers such that $(nb_n^2)/[\log(n)a_n] \to \infty$ and $a_n \to 0$. Then

$$\sup_{\{\sigma_n : b_n \leq \sigma_n \leq a_n\}} |\hat{\delta}_{n,\sigma_n} - \delta| \to 0 \quad a.s. \tag{S1.5}$$

Define

$$f_{n,\sigma_n}(t) = \frac{1}{\sigma_n}\int_{-\infty}^{\infty} K\left(\frac{t - y}{\sigma_n}\right) f(y)dy.$$

Choose r such that $0 < r < r_0$, where r_0 is obtained by Proposition 5.1 (Romano,

1988) so that

$$\lim_{n\to\infty} \sup_{\{\sigma_n : 0<\sigma_n\le a_n\}} \sup_{\{t:|t-\delta|<r_0\}} |f_{n,\sigma_n}(t) - f(t)| = 0.$$

Then by Corollary 5.1 of Romano (1988), we have

$$\sup_{\{\sigma_n : b_n\le\sigma_n\le a_n\}} \sup_{t} |\hat{f}_{n,\sigma_n}(t) - f_{n,\sigma_n}(t)| \to 0 \quad a.s.$$

Combining the above equations yields

$$\sup_{\{\sigma_n : b_n\le\sigma_n\le a_n\}} \sup_{\{t:|t-\delta|<r\}} |\hat{f}_{n,\sigma_n}(t) - f(t)| \to 0 \quad a.s. \tag{S1.6}$$

The assumptions imply for any $r > 0$,

$$\limsup_{n\to\infty} \sup_{\{\sigma_n : b_n\le\sigma_n\le a_n\}} \sup_{\{t:|t-\delta|>r\}} f_{n,\sigma_n}(t) < f(\delta). \tag{S1.7}$$

This yields

$$\limsup_{n\to\infty} \sup_{\{\sigma_n : b_n\le\sigma_n\le a_n\}} \sup_{\{t:|t-\delta|\ge r\}} \hat{f}_{n,\sigma_n}(t) < f(\delta) \quad a.s. \tag{S1.8}$$

It follows from (S1.6) and (S1.8) that $P(\sup_{\{\sigma_n : b_n\le\sigma_n\le a_n\}} |\hat{\delta}_{n,\sigma_n} - \delta| < r) = 1$. This proves (S1.5). Hence by (S1.5)

$$\limsup_{n\to\infty} |\hat{\delta}_{n,\sigma_n} - \delta| \le \limsup_{n\to\infty} \sup_{\{\sigma_n : b_n\le\sigma_n\le a_n\}} |\hat{\delta}_{n,\sigma_n} - \delta| \to 0 \quad a.s.$$

It follows that $\sup_{\{\sigma_n : b_n\le\sigma_n\le a_n\}} \sup_t \hat{f}_{n,\sigma_n}(t) \to f(\delta)$. □

Lemma S5. *Under the assumptions (A1)-(A4) , the following equation holds, for all* $\epsilon > 0$,

$$\lim_{n\to\infty} P\left\{ \sup |\Gamma_n(\boldsymbol{\sigma}) - \Gamma(\boldsymbol{\sigma})| > \epsilon \right\} = 0. \tag{S1.9}$$

Proof. Proof of Lemma S6. Suppose equation (S1.9) is wrong. This means there exist $\epsilon > 0$, $\alpha > 0$ and a subsequence M of positive integers and $(\boldsymbol{\sigma}_m^*)_{m \in M}$ such that

$$\text{for all } m \in M, P\{|\Gamma_m(\boldsymbol{\sigma}_m^*) - \Gamma(\boldsymbol{\sigma}_m^*)| > \epsilon\} \geq \alpha. \tag{S1.10}$$

Since $(\boldsymbol{\sigma}_m^*) \to 0$ is bounded and at least a subsequence has a limit, there exists $\boldsymbol{\sigma}^* = \lim_{m\to\infty} \boldsymbol{\sigma}_m^*$.

Consider the following inequality,

$$|\Gamma_m(\boldsymbol{\sigma}_m^*) - \Gamma(\boldsymbol{\sigma}_m^*)| \leq |\Gamma_m(\boldsymbol{\sigma}_m^*) - \Gamma_m(\boldsymbol{\sigma}^*)| + |\Gamma_m(\boldsymbol{\sigma}^*) - \Gamma(\boldsymbol{\sigma}^*)| + |\Gamma(\boldsymbol{\sigma}^*) - \Gamma(\boldsymbol{\sigma}_m^*)|. \tag{S1.11}$$

The third term of the right side of (S1.11) converges to zero due to the continuity of Γ. To handle with the second term, we define a version of Γ_n using the true modes instead of the empirical ones:

$$\Gamma_n^*(\boldsymbol{\sigma}) = \frac{1}{n}\left[\sum_{i:\mathcal{C}_i=1} I\{\lambda(\boldsymbol{z},\boldsymbol{\sigma}) > 0\} + \sum_{i:\mathcal{C}_i=2} I\{\lambda(\boldsymbol{z},\boldsymbol{\sigma}) \leq 0\}\right],$$

where $\lambda(Z,\boldsymbol{\sigma}) = \sum_{j=1}^p \{\Psi_{1j}(Z_j,\sigma_{1jn}) - \Psi_{2j}(Z_j,\sigma_{2jn})\} > 0$. Thus we have

$$|\Gamma_m(\boldsymbol{\sigma}) - \Gamma(\boldsymbol{\sigma}^*)| \leq |\Gamma_m(\boldsymbol{\sigma}^*) - \Gamma_m^*(\boldsymbol{\sigma}^*)| + |\Gamma_m^*(\boldsymbol{\sigma}^*) - \Gamma(\boldsymbol{\sigma}^*)|.$$

It follows that $\lim_{m\to\infty} |\Gamma_m^*(\boldsymbol{\sigma}^*) - \Gamma(\boldsymbol{\sigma}^*)| = 0$ almost surely due to the strong law of large numbers. In addition, for given $\boldsymbol{z}$ and $\sigma_{1jn}, \sigma_{2jn}$, Ψ_{ij} and Ψ_{2j} are continuous in δ_{1j} and δ_{2j}, respectively. Besides, based on Lemma (S4), modes are strongly consistent given the selected bandwidth, thus $\lim_{m\to\infty} |\Gamma_m(\boldsymbol{\sigma}^*) - \Gamma_m^*(\boldsymbol{\sigma}^*)| = 0$ almost surely.

We next consider the first term of the right side of (S1.11). Since we have

$$\begin{aligned}|\delta_{2jm}(\sigma_2^*) - \delta_{2jm}(\sigma_2)| &\leq |\delta_{2jm}(\sigma_2^*) - \delta_{2j}(\sigma_2^*)| + |\delta_{2j}(\sigma_2^*) - \delta_{2j}(\sigma_2)| \\ &+ |\delta_{2j}(\sigma_2) - \delta_{2jm}(\sigma_2)|. \end{aligned} \tag{S1.12}$$

and

$$\begin{aligned}|\delta_{1jm}(\sigma_1^*) - \delta_{1jm}(\sigma_1)| &\leq |\delta_{1jm}(\sigma_1^*) - \delta_{1j}(\sigma_1^*)| + |\delta_{1j}(\sigma_1^*) - \delta_{1j}(\sigma_1)| \\ &+ |\delta_{1j}(\sigma_1) - \delta_{1jm}(\sigma_1)|. \end{aligned} \tag{S1.13}$$

From Theorem 1.1 of Romano (1988), under assumption (A2) and (A3), we obtain $\lim_{n\to\infty} |\delta_{ij}(\sigma_i) - \delta_{ijm}(\sigma_i)| = 0$ almost surely for $i = 1, 2$. This makes the first and third terms of the right side of (S1.12) converge to zero almost surely. The second term converges to zero because of assumption (A4'). Thereby, for $m \to \infty$,

$$|\delta_{ijm}(\sigma_2^*) - \delta_{ijm}(\sigma_2)| \to 0, \quad a.s. \text{ for } \quad i = 1, 2. \tag{S1.14}$$

Let $\Delta_n(\sigma_1, \sigma_2, z) = \sum_{j=1}^p \{\Psi_{1jn}(z, \sigma_1) - \Psi_{2jn}(z, \sigma_2)\}$, $\Delta(\sigma_1, \sigma_2, z) = \sum_{j=1}^p \{\Psi_{1j}(z, \sigma_1) - \Psi_{2j}(z, \sigma_2)\}$. For $\epsilon > 0$, define

$$Z_\epsilon = \{z : |\Delta(\sigma_1^*, \sigma_2^*, z)| > \epsilon\} \cap \left(z : \sum_{j=1}^p |z_j| \leq 1/\epsilon\right).$$

Thus, we have

$$|\Gamma_m(\boldsymbol{\sigma}_m^*) - \Gamma_m(\boldsymbol{\sigma}^*)| = \frac{1}{m}\Big(\sum_{i:C_i=1,Z_i\notin Z_\epsilon} [I\{\Delta_m(\sigma_{1m}^*, \sigma_{2m}^*, Z_i) > 0\} - I\{\Delta_m(\sigma_1^*, \sigma_2^*, Z_i) > 0\}] + \sum_{i:C_i=2,Z_i\notin Z_\epsilon} [I\{\Delta_m(\sigma_{1m}^*, \sigma_{2m}^*, Z_i) \le 0\} - I\{\Delta_m(\sigma_1^*, \sigma_2^*, Z_i) \le 0\}] + \sum_{i:C_i=1,Z_i\in Z_\epsilon} [I\{\Delta_m(\sigma_{1m}^*, \sigma_{2m}^*, Z_i) > 0\} - I\{\Delta_m(\sigma_1^*, \sigma_2^*, Z_i) > 0\}] + \sum_{i:C_i=2,Z_i\in Z_\epsilon} [I\{\Delta_m(\sigma_{1m}^*, \sigma_{2m}^*, Z_i) \le 0\} - I\{\Delta_m(\sigma_1^*, \sigma_2^*, Z_i) \le 0\}]\Big).$$

For large m and arbitrarily small $\eta > 0$,

$$\frac{1}{m}\Big|\Big(\sum_{i:C_i=1,Z_i\notin Z_\epsilon} [I\{\Delta_m(\sigma_{1m}^*, \sigma_{2m}^*, Z_i) > 0\} - I\{\Delta_m(\sigma_1^*, \sigma_2^*, Z_i) > 0\}] + \sum_{i:C_i=2,Z_i\notin Z_\epsilon} [I\{\Delta_m(\sigma_{1m}^*, \sigma_{2m}^*, Z_i) \le 0\} - I\{\Delta_m(\sigma_1^*, \sigma_2^*, Z_i) \le 0\}]\Big)\Big| \le 1 - P(Z_\epsilon) + \eta, \quad a.s.$$

In addition, we have $|\Delta_m(\sigma_{1m}^*, \sigma_{2m}^*, Z_i) - \Delta_m(\sigma_1^*, \sigma_2^*, Z_i)| \le |\sum_{j=1}^p \{\Psi_{1jm}(Z_i, \sigma_{1m}^*) - \Psi_{1jm}(Z_i, \sigma_1^*)\}| + |\sum_{j=1}^p \{\Psi_{2jm}(Z_i, \sigma_{2m}^*) - \Psi_{2jm}(Z_i, \sigma_2^*)\}|$, then by (S1.2) and (S1.3),

$$\begin{aligned} & |\Delta_m(\sigma_{1m}^*, \sigma_{2m}^*, Z_i) - \Delta_m(\sigma_1^*, \sigma_2^*, Z_i)| \\ \le\ & \sum_{j=1}^p c_1|z_j - \delta_{1j}||\frac{1}{\sigma_{1m}^*} - \frac{1}{\sigma_1^*}| + c_0|\frac{1}{\sigma_{1m}^*} - \frac{1}{\sigma_1^*}| + \sum_{j=1}^p c_1|z_j - \delta_{2j}||\frac{1}{\sigma_{2m}^*} - \frac{1}{\sigma_2^*}| + c_0|\frac{1}{\sigma_{2m}^*} - \frac{1}{\sigma_2^*}|. \end{aligned}$$

Since $|\sigma_{1m}^* - \sigma_1^*| \to 0$ and $|\sigma_{2m}^* - \sigma_2^*| \to 0$, by (S1.19),and in addition by $\sum_{j=1}^p |Z_j| < 1/\epsilon$ for $Z_j \in Z_\epsilon$, $|\Delta_m(\sigma_{1m}^*, \sigma_{2m}^*, Z_i) - \Delta_m(\sigma_1^*, \sigma_2^*, Z_i)|$ is arbitrarily small for large enough m. If $\epsilon \to 0$, we have $P(Z_\epsilon) \to 1$ by assumption (A4). Thus, for large m, $|\Gamma_m(\sigma_m^*) - \Gamma_m(\sigma^*)| \to 0$ a.s. This is in contradiction to equation (S1.16), which completes the proof. □

Lemma S6. *Under the assumptions (A1)-(A3) and (A4') , the following equation holds, for all* $\epsilon > 0$,

$$\lim_{n\to\infty} P\left\{ \sup_{\theta\in T} |\Gamma_n(\theta) - \Gamma(\theta)| > \epsilon \right\} = 0. \quad \text{(S1.15)}$$

Proof. Proof of Lemma S6. Suppose equation (S1.15) is wrong. This means there exist $\epsilon > 0$, $\alpha > 0$ and a subsequence M of positive integers and $(\theta^*_m)_{m\in M}$ such that

$$\text{for all } m \in M, P\{|\Gamma_m(\theta^*_m) - \Gamma(\theta^*_m)| > \epsilon\} \geq \alpha. \quad \text{(S1.16)}$$

Since $(\theta_m)_{m\in M} \in T^M$ is bounded and at least a subsequence has a limit, there exists $\theta^* = \lim_{m\to\infty} \theta^*_m$.

Consider the following inequality,

$$|\Gamma_m(\theta^*_m) - \Gamma(\theta^*_m)| \leq |\Gamma_m(\theta^*_m) - \Gamma_m(\theta^*)| + |\Gamma_m(\theta^*) - \Gamma(\theta^*)| + |\Gamma(\theta^*) - \Gamma(\theta^*_m)|. \quad \text{(S1.17)}$$

The third term of the right side of (S1.17) converges to zero due to the continuity of Γ. To handle with the second term, we define a version of Γ_n using the true modes instead of the empirical ones:

$$\Gamma^*_n(\theta) = \frac{1}{n}\left[\sum_{i:\mathcal{C}_i=1} I\{\lambda(\boldsymbol{z}, \theta, \sigma_{1j}) > 0\} + \sum_{i:\mathcal{C}_i=2} I\{\lambda(\boldsymbol{z}, \theta, \sigma_{1j}) \leq 0\} \right],$$

where $\lambda(Z, \theta, \sigma_{1j}) = \sum_{j=1}^{p}\{\Psi_{1j}(Z_j, \sigma_{1jn}) - \Psi_{2j}(Z_j, \theta\sigma_{1j})\} > 0$. Thus we have

$$|\Gamma_m(\theta^*) - \Gamma(\theta^*)| \leq |\Gamma_m(\theta^*) - \Gamma^*_m(\theta^*)| + |\Gamma^*_m(\theta^*) - \Gamma(\theta^*)|.$$

It follows that $\lim_{m\to\infty} |\Gamma^*_m(\theta^*) - \Gamma(\theta^*)| = 0$ almost surely due to the strong law of large numbers. In addition, for given $\boldsymbol{z}$ and σ_1, θ, Ψ_{ij} and Ψ_{2j} are continuous in δ_{1j}

and δ_{2j}, respectively. Besides, modes are strongly consistent, thus $\lim_{m\to\infty} |\Gamma_m(\theta^*) - \Gamma_m^*(\theta^*)| = 0$ almost surely according to equation (S1.1).

We next consider the first term of the right side of (S1.17). Since we have

$$\begin{aligned} |\delta_{2jm}(\theta_m^*\sigma_1) - \delta_{2jm}(\theta^*\sigma_1)| &\leq |\delta_{2jm}(\theta^*\sigma_1) - \delta_{2j}(\theta^*\sigma_1)| + |\delta_{2jm}(\theta^*\sigma_1) - \delta_{2j}(\theta_m^*\sigma_1)| \\ &+ |\delta_{2j}(\theta_m^*\sigma_1) - \delta_{2j}(\theta^*\sigma_1)|. \end{aligned} \tag{S1.18}$$

From Theorem 1.1 (Romano,1988), under assumption (A2) and (A3), we obtain $\lim_{n\to\infty} \sup_{\theta\in T} |\delta_{2j}(\theta\sigma_1) - \delta_{2jm}(\theta\sigma_1)| = 0$ almost surely. This makes the first two term of the right side of (S1.18) converge to zero almost surely. The last term converges to zero because of assumption A4. Thereby, for $m \to \infty$,

$$|\delta_{2jm}(\theta_m^*\sigma_1) - \delta_{2jm}(\theta^*\sigma_1)| \to 0, a.s. \tag{S1.19}$$

Let $\Delta_n(\theta, \sigma_1, z) = \sum_{i=1}^p \{\Psi_{1jn}(z, \sigma_1) - \Psi_{2jn}(z, \sigma_1\theta)\}$, $\Delta(\theta, \sigma_1, z) = \sum_{i=1}^p \{\Psi_{1j}(z, \sigma_1) - \Psi_{2j}(z, \sigma_1\theta)\}$. For $\epsilon > 0$, define

$$Z_\epsilon = \{z : |\Delta(\theta^*, \sigma_1, z)| > \epsilon\} \cap \left(z : \sum_{j=1}^p |z_j| \leq 1/\epsilon\right).$$

Thus, we have

$$\begin{aligned} |\Gamma_m(\theta_m^*) - \Gamma_m(\theta^*)| = \frac{1}{m}\Bigg(& \sum_{i:C_i=1, Z_i\notin Z_\epsilon} [I\{\Delta_m(\theta_m^*, \sigma_1, Z_i) > 0\} - I\{\Delta_m(\theta^*, \sigma_1, Z_i) > 0\}] + \\ & \sum_{i:C_i=2, Z_i\notin Z_\epsilon} [I\{\Delta_m(\theta_m^*, \sigma_1, Z_i) \leq 0\} - I\{\Delta_m(\theta^*, \sigma_1, Z_i) \leq 0\}] + \\ & \sum_{i:C_i=1, Z_i\in Z_\epsilon} [I\{\Delta_m(\theta_m^*, \sigma_1, Z_i) > 0\} - I\{\Delta_m(\theta^*, \sigma_1, Z_i) > 0\}] + \\ & \sum_{i:C_i=2, Z_i\in Z_\epsilon} [I\{\Delta_m(\theta_m^*, \sigma_1, Z_i) \leq 0\} - I\{\Delta_m(\theta^*, \sigma_1, Z_i) \leq 0\}]\Bigg). \end{aligned}$$

For large m and arbitrarily small $\eta > 0$,

$$\frac{1}{m}\Bigg|\Bigg(\sum_{i:C_i=1,Z_i\notin Z_\epsilon}[I\{\Delta_m(\theta_m^*,\sigma_1,Z_i)>0\}-I\{\Delta_m(\theta^*,\sigma_1,Z_i)>0\}]+ \sum_{i:C_i=2,Z_i\notin Z_\epsilon}[I\{\Delta_m(\theta_m^*,\sigma_1,Z_i)\le 0\}-I\{\Delta_m(\theta^*,\sigma_1,Z_i)\le 0\}]\Bigg)\Bigg| \le 1-P(Z_\epsilon)+\eta, \quad a.s.$$

In addition, we have $|\Delta_m(\theta_m^*,Z_i)-\Delta_m(\theta^*,Z_i)| = \sum_{j=1}^p\{\Psi_{2jm}(Z_i,\theta^*\sigma_{1j})-\Psi_{2jm}(Z_i,\theta_m^*\sigma_{2j})\}$, then by (S1.4)

$$|\Delta_m(\theta_m^*,Z_i)-\Delta_m(\theta^*,Z_i)| \le \sum_{j=1}^p\frac{1}{\sigma_{1j}}\left\{2c_0\left|\frac{1}{\theta^*}-\frac{1}{\theta_m^*}\right|+\frac{c_1|z_j|}{\sigma_{1j}\theta^*}\left|\frac{1}{\theta^*}-\frac{1}{\theta_m^*}\right|+\frac{c_1}{\theta^*}\left|\frac{\delta_{2j}(\theta^*\sigma_{1j})}{\theta^*\sigma_{1j}}-\frac{\delta_{2j}(\theta_m^*\sigma_{1j})}{\theta_m^*\sigma_{1j}}\right|\right\}.$$

Since $|\theta_m^*-\theta^*|\to 0$, by (S1.19), we have $|\delta_{2jm}(\theta_m^*\sigma_1)-\delta_{2jm}(\theta^*\sigma_1)|\to 0$, a.s. and $\left|\frac{1}{\theta^*}-\frac{1}{\theta_m^*}\right|\to 0$ a.s. In addition by $\sum_{j=1}^p|Z_j|<1/\epsilon$ for $Z_j\in Z_\epsilon$, $|\Delta_m(\theta_m^*,Z_i)-\Delta_m(\theta^*,Z_i)|$ is arbitrarily small for large enough m, and it follows that $\Delta_m(\theta_m^*,Z_i)$ and $\Delta_m(\theta^*,Z_i)$ will be either both positive and both negative for sufficiently large m, if $Z_i\in Z_\epsilon$. Thus the corresponding indicator functions would be the same, almost surely. If $\epsilon\to 0$, we have $P(Z_\epsilon)\to 1$ by assumption (A5). Thus, for large m, $|\Gamma_m(\theta_m^*)-\Gamma_m(\theta^*)|\to 0$ a.s. This is in contradiction to equation (S1.16), which completes the proof. □

Proof. Proof of Theorem 1.

$$\begin{aligned} & P(|\Gamma(\boldsymbol{\sigma})-\Gamma(\hat{\boldsymbol{\sigma}}_n)|>\epsilon) \\ \le\ & P(|\Gamma(\boldsymbol{\sigma})-\Gamma_n(\boldsymbol{\sigma})|>\epsilon/3)+P(|\Gamma_n(\boldsymbol{\sigma})-\Gamma_n(\hat{\boldsymbol{\sigma}}_n)|>\epsilon/3)+P(|\Gamma_n(\hat{\boldsymbol{\sigma}}_n)-\Gamma(\hat{\boldsymbol{\sigma}}_n)|>\epsilon/3) \\ =\ & A_1+A_2+A_3. \end{aligned}$$

According to Lemma S5, we have $A_1 \to 0$ and $A_3 \to 0$. Since by lemma S2, $\Gamma_n(\hat{\boldsymbol{\sigma}}_n) \to \Gamma_n(\boldsymbol{\sigma})$, we complete the proof.

Proof. Proof of Theorem 2.

First, we will show that Γ is a continuous function of θ. By assumption (A4), we have for given z, σ_{1j}, $j = 1, \ldots, p$, Ψ_{2j} is a continuous function of θ, consequently $\lambda(z, \theta, \boldsymbol{\sigma}_1)$ is a continuous function of θ. The fact that $|I(\lambda(z, \theta, \sigma_1) > 0)| \leq 1$ and $|I(\lambda(z, \theta, \sigma_1) \leq 0)| \leq 1$, together with the dominated convergence theorem, implies the continuity of Γ.

Then we prove Theorem 1. Note that

$$
\begin{aligned}
& P(|\Gamma(\tilde{\theta}) - \Gamma(\hat{\theta}_n)| > \epsilon) \\
\leq \ & P(|\Gamma(\tilde{\theta}) - \Gamma_n(\tilde{\theta})| > \epsilon/3) + P(|\Gamma_n(\tilde{\theta}) - \Gamma_n(\hat{\theta}_n)| > \epsilon/3) + P(|\Gamma_n(\hat{\theta}_n) - \Gamma(\hat{\theta}_n)| > \epsilon/3) \\
= \ & E_1 + E_2 + E_3.
\end{aligned}
$$

According to Lemma S6, we have $E_1 \to 0$ and $E_3 \to 0$. Next we deal with item E_2. Again with Lemma S6, we have for large n, there exist $\epsilon > 0$ such that $|\Gamma(\tilde{\theta}) - \Gamma_n(\tilde{\theta})| < \epsilon/6$ and $|\Gamma_n(\hat{\theta}_n) - \Gamma(\hat{\theta}_n)| < \epsilon/6$. In addition, by definition, $\Gamma_n(\hat{\theta}_n) \geq \Gamma_n(\tilde{\theta})$ and $\Gamma(\tilde{\theta}) \geq \Gamma(\hat{\theta}_n)$. Thus we have $0 \leq \Gamma(\tilde{\theta}) - \Gamma(\hat{\theta}_n) \leq \Gamma_n(\tilde{\theta}) - \Gamma_n(\hat{\theta}_n) + \epsilon/3$. It follows that $0 \leq \Gamma_n(\hat{\theta}_n) - \Gamma_n(\tilde{\theta}) \leq \epsilon/3$. Therefore $E_2 \to 0$. This completes the proof. □

□

Proof. Proof of Theorem 3. Let ν_{Z_k} denote the mode of Z_k, and let $\nu_{XZ_k} = \nu_{X_k} - \nu_{Z_k}$ and $\nu_{YZ_k} = \nu_{X_k} - \nu_{Z_k}$. Write $\mathcal{U}_k$ and $\mathcal{V}_k$ for samples of sizes m and n, respectively, of independent random variables each distributed as U_k and all independent of Z_k. Let $W_k = Z_k - \nu_{Z_k}$ be another random variable distributed as U_k. Assumption (A8) together with both m and n diverge as $p \to \infty$ imply that

$$\text{for each } \epsilon > 0, \quad \lim_{p\to\infty} \sup_{1\le k\le p} \left\{ P\left(1-\gamma K\left(\frac{\text{mode}(\mathcal{U}_k)}{\sigma_k}\right) > \epsilon\right) + P\left(1-\gamma K\left(\frac{\text{mode}(\mathcal{V}_k)}{\sigma_k}\right) > \epsilon\right)\right\} = 0.$$

This property and the fact that $\text{mode}(\mathcal{X}_k) - \nu_{X_k}$ and $\text{mode}(\mathcal{Y}_k) - \nu_{Y_k}$ are distributed as $\text{mode}(\mathcal{U}_k)$ and $\text{mode}(\mathcal{V}_k)$, respectively, imply that for each $\epsilon > 0$ and with P denoting either P_X or P_Y,

$$\begin{aligned}\lim_{p\to\infty} \sup_{1\le k\le p} \Bigg[\quad & P\left(1-\gamma K\left(\frac{(\text{mode}(\mathcal{X}_k) - Z_k) - (\nu_{XZ_k} - W_k)}{\sigma_k}\right) > \epsilon\right) \\ + \quad & P\left(1-\gamma K\left(\frac{(\text{mode}(\mathcal{Y}_k) - Z_k) - (\nu_{YZ_k} - W_k)}{\sigma_k}\right) > \epsilon\right)\Bigg] = 0.\end{aligned}$$

Therefore, define

$$\begin{aligned} D_k &= \gamma\left| K\left(\frac{\text{mode}(\mathcal{X}_k) - Z_k}{\sigma_k}\right) - K\left(\frac{\text{mode}(\mathcal{Y}_k) - Z_k}{\sigma_k}\right) - \left(K\left(\frac{\nu_{XZ_k} - W_k}{\sigma_k}\right) - K\left(\frac{\nu_{YZ_k} - W_k}{\sigma_k}\right)\right)\right| \\ &= \big|\Psi_k(Z_k - \text{mode}(\mathcal{X}_k), \sigma_k) - \Psi_k(Z_k - \text{mode}(\mathcal{Y}_k), \sigma_k) \\ &\qquad -[\Psi_k(W_k - \nu_{XZ_k}, \sigma_k) - \Psi_k(W_k - \nu_{YZ_k}, \sigma_k)]\big|,\end{aligned}$$

where $\Psi_k(x, y) = \gamma K\{x/y\}$, $\gamma = 1/K(0)$ and take $\mathcal{A}_k(\epsilon)$ to be the set on which $D_k > \epsilon$, we have for either choice of P,

$$\text{for each } \epsilon > 0, \quad \lim_{p\to\infty} \sup_{1\le k\le p} P\{\mathcal{A}_k(\epsilon)\} = 0. \tag{S1.20}$$

Let

$$D(Z) \equiv \Big| \sum_{k=1}^{p} [\Psi_k(Z_k - \mathrm{mode}(\mathcal{X}_k), \sigma_k) - \Psi_k(Z_k - \mathrm{mode}(\mathcal{Y}_k), \sigma_k)] - \sum_{k=1}^{p} [\Psi_k(W_k - \nu_{XZ_k}, \sigma_k) - \Psi_k(W_k - \nu_{YZ_k}, \sigma_k)] \Big|,$$

thus we have $D(Z) \leq p\epsilon + \sum_{k=1}^{p} D_k I\{\mathcal{A}_k(\epsilon)\}$. Because of the assumption (A6), it follows that

$$\lim_{\lambda\to\infty} \sup_{k\geq 1} E\big\{D_k I(D_k > \lambda)\big\} = 0. \tag{S1.21}$$

Equation (S1.20) together with (S1.21) imply that

$$\text{for each } \epsilon > 0, \quad \sum_{k=1}^{p} E[D_k I\{\mathcal{A}_k(\epsilon)\}] = o(p).$$

Since $D(Z) \leq p\epsilon + \sum_{k=1}^{p} D_k I\{\mathcal{A}_k(\epsilon)\}$ holds for each $\epsilon > 0$,

$$D(Z) = o_p(p).$$

In addition by assumption (A10), for each $\epsilon > 0$ we can choose $\lambda > 0$ so large that for all p,

$$\sum_{k=1}^{p} E\bigg\{\Big|\Psi_k(W_k - \nu_{XZ_k}, \sigma_k) - \Psi_k(W_k - \nu_{YZ_k}, \sigma_k)\Big| I(|W_k| > \lambda)\bigg\} \leq p\epsilon, \tag{S1.22}$$

since $\Big|\Psi_k(W_k - \nu_{XZ_k}, \sigma_k) - \Psi_k(W_k - \nu_{YZ_k}, \sigma_k)\Big| \leq (c_1\gamma/\sigma_k)|\nu_{XZ_k} - \nu_{YZ_k}|$ is uniformly bounded. Define

$$S_\lambda = \sum_{k=1}^{p} \Big(\Psi_k(W_k - \nu_{XZ_k}, \sigma_k) - \Psi_k(W_k - \nu_{YZ_k}, \sigma_k)\Big) I\{|W_k| \leq \lambda\}$$

and $d(Z) = S_\infty$. (A9) and (A10) imply that for each fixed λ, $var(S_\lambda) = o_p(p^2)$ as $p \to \infty$, thereby it follows that $S_\lambda - E(S_\lambda) = o_p(p)$. Since ϵ in (S1.22) can be chosen arbitrarily small, thus we have

$$d(Z) = E\{d(Z)\} + o_p(p).$$

By the definition of D_k, it is easy to derive that

$$\sum_{k=1}^{p} \Big\{ \Psi_k\big(Z_k - \text{mode}(\mathcal{X}_k), \sigma_k\big) - \Psi_k\big(Z_k - \text{mode}(\mathcal{Y}_k), \sigma_k\big) \Big\} = E(d(Z)) + o_p(p). \quad \text{(S1.23)}$$

Given $\epsilon > 0$, let $\mathcal{B}_\epsilon = \mathcal{B}_\epsilon(p)$ denote the set of indices $k \in \{1, \dots, p\}$ such that $|\nu_{X_k} - \nu_{Y_k}| \geq (\sigma_k/(c_1\gamma))\epsilon$. If $k \in \{1, \dots, p\}$ but $k \notin \mathcal{B}_\epsilon$, then by assumption (A1)

$$|\Psi_k(\nu_{XZ_k}, W_k) - \Psi_k(\nu_{YZ_k}, W_k)| \leq (c_1\gamma)|\nu_{X_k} - \nu_{Y_k}|/\sigma_k \leq \epsilon.$$

Define $d_k = E\{\Psi_k(\nu_{XZ_k}, W_k) - \Psi_k(\nu_{YZ_k}, W_k)\}$, we have

$$\Big| E(d(Z)) - \sum_{k \in \mathcal{B}_\epsilon} d_k \Big| \leq p\epsilon. \quad \text{(S1.24)}$$

Given $c > 0$, let $a(c)$ equal the infimum of $E\Psi_k(U_k, \sigma_k) - E\Psi_k(U_k + x, \sigma_{kx})$ over all $|x| > c$ and all $k \geq 1$. Due to assumption (A7), $a(c) > 0$. If Z is from the X population, then $\nu_{XZ_k} = 0$ and $\nu_{YZ_k} = \nu_{Y_k} - \nu_{Z_k} := -\delta_k$, thus for $k \in \mathcal{B}_c$,

$$d_k = E_X\{\Psi_k(\nu_{XZ_k}, W_k) - \Psi_k(\nu_{YZ_k}, W_k)\} \geq a(c)I(|\delta_k| > c).$$

Here E_X denotes expectation under the assumption that Z is drawn from $\mathcal{F}_X$. Thereby, if $c \geq \epsilon$,

$$\sum_{k \in \mathcal{B}_\epsilon} d_k \geq a(c)|\mathcal{B}_c|, \quad \text{(S1.25)}$$

here $|\mathcal{B}_c|$ denotes the size of $\mathcal{B}_c$.

(S1.24) and (S1.25) imply that

$$E_X(d(Z)) \geq \sum_{k \in \mathcal{B}_\epsilon} d_k - p\epsilon \geq a(c)|\mathcal{B}_c| - p\epsilon. \tag{S1.26}$$

In addition by assumption (A10), the proportion of $k \in \{1, \ldots, p\}$ for which $|\nu_{X_k} - ny| > \epsilon$ is bounded away from zero as p diverges, if c is taken sufficiently small, then there exists a constant $\alpha(c) > 0$ such that $|\mathcal{B}_c| \geq \alpha(c)p$ for all large p. Thus (S1.26) becomes

$$E_X(d(Z)) \geq a(c)\alpha(c)p - p\epsilon,$$

and $p^{-1}E_X(d(Z)) \geq a(c)\alpha(c) - \epsilon$, since both $a(c)$ and $\alpha(c)$ are positive and ϵ can be chosen arbitrarily small, thus we have

$$\liminf_{p\to\infty} p^{-1}E_X(d(Z)) > 0.$$

Again by equation (S1.23), there exists a constant $b > 0$ such that

$$\liminf_{p\to\infty} P_X\left\{ \sum_{k=1}^{p} p^{-1}\{\Psi_k(Z_k - \text{mode}(\mathcal{X}_k), \sigma_k) - \Psi_k(Z_k - \text{mode}(\mathcal{Y}_k), \sigma_k)\} \geq b \right\} = 1,$$

that is

$$\liminf_{p\to\infty} P_X\left\{ \sum_{k=1}^{p} \{\Psi_k(Z_k - \text{mode}(\mathcal{X}_k), \sigma_k) - \Psi_k(Z_k - \text{mode}(\mathcal{Y}_k), \sigma_k)\} \geq bp \right\} = 1,$$

this means if Z is from population $\mathcal{F}_X$, the classifier assigns Z to the X population, converges to 1. The conclusion also holds if Z is from $\mathcal{F}_Y$.

The proof for empirically optimal unimodal classifier is similar. After definitions have been adapted to empirically optimal unimodal classifier, the above-mentioned statements hold uniformly for $\theta \in T$. This completes the proof. $\square$

S2 Rate of convergence of the unimodal classifier

In applications, it is desirable to construct a data-driven classification rule based on two observed random samples, $\mathbf{X}_1^{(1)}, \ldots, \mathbf{X}_m^{(1)} \overset{iid}{\sim} \mathcal{F}_1$ and $\mathbf{Y}_1^{(2)}, \ldots, \mathbf{Y}_n^{(2)} \overset{iid}{\sim} \mathcal{F}_2$. Consider a binary classification problem. The data are given as $(\mathbf{X}_1, 1), \ldots, (\mathbf{X}_m, 1)$, $(\mathbf{Y}_1, 2), \ldots, (\mathbf{Y}_n, 2)$, where $\mathbf{X}_i \in \mathcal{X} \subset \mathbb{R}^p$ and $\mathbf{Y}_i \in \mathcal{Y} \subset \mathbb{R}^p$ are input vectors, $\{1, 2\}$ are class labels. The goal is to find a function $\hat{\mathcal{M}} : \mathbb{R}^p \to \{1, 2\}$ that minimizes the probability of error. Given the observed data, the performance of classification rule is measured by the misclassification error, $R(\hat{\mathcal{M}}) = P(\mathcal{M}(\mathbf{Z}) \neq \mathit{label}(\mathbf{Z}))$, and $\mathbf{Z}$ is independent of observed $(\mathbf{X}, \mathbf{Y})$s, $\mathit{label}(\mathbf{Z})$ denotes the true class of $\mathbf{Z}$. The minimum possible probability of error is the Bayes risk, denoted by

$$R^* = \inf_{\mathcal{M}} R(\mathcal{M}) = E_{\mathbf{X}}[\min(\eta(\mathbf{Z}), 1 - \eta(\mathbf{Z}))],$$

where the infimum is taken over all measurable classifiers, and $\eta(\mathbf{z}) = P(C = 1 | \mathbf{Z} = \mathbf{z})$ denotes the posterior probability function or the regression function of C on $\mathbf{Z} = \mathbf{z}$. The infimum is achieved by the Bayes classifier $\mathcal{M}^*(\mathbf{z}) = I(\eta(\mathbf{z}) \geq 1/2) + 1$, where I denotes the indicator function, that is $\mathcal{M}^*(\mathbf{z}) = I(\pi_1 f_1(\mathbf{z}) - \pi_2 f_2(\mathbf{z}) \geq 0) + 1$, where

$f_1(\cdot)$ and $f_2(\cdot)$ are probability density functions for $\mathcal{F}_1$ and $\mathcal{F}_2$, respectively. Based on the label samples, one can estimate the regression function $\eta(\mathbf{x})$ or estimate the probability density functions f_1 and f_2, and the corresponding classification rule is $\widetilde{\mathcal{M}}(\mathbf{z}) = I(\tilde{\eta}(\mathbf{z}) - 1/2 \geq 0) + 1$ or $\widetilde{\mathcal{M}}(\mathbf{z}) = I(\pi_1 \tilde{f}_1(\mathbf{z}) - \pi_2 \tilde{f}_2(\mathbf{z}) \geq 0) + 1$.

The accuracy of an empirical decision rule $\widetilde{\mathcal{M}}_n$ can be characterized by excess risk,

$$\mathcal{E}(\widetilde{\mathcal{M}}_n) = ER(\widetilde{\mathcal{M}}_n) - R(\mathcal{M}^*) = E\big(|2\eta(\mathbf{Z}) - 1| I\{\widetilde{\mathcal{M}}_n(\mathbf{Z}) \neq \mathcal{M}^*(\mathbf{Z})\}\big)$$

where E denotes expectation. A key problem in classification is to construct classifiers with small excess risk. Optimal classifiers can be defined as those having the best possible rate of convergence of $\mathcal{E}(\widetilde{\mathcal{M}}_n) \to 0$ as $n \to \infty$. Of course, this rate, and thus the optimal classifier, depend on the assumptions on the joint distribution of $(\mathbf{X}, \mathbf{Y})$.

In the sequel our focus lies on the rate of convergence of the excess error probability $E\{\widehat{\mathcal{M}}_n\} - \mathcal{M}^*$, where $\widehat{\mathcal{M}}_n$ is the unimodal classification rule defined as follows. The unimodal classifier $\widehat{\mathcal{M}}_n$ is to classify a future data point $\boldsymbol{z}$ drawn from these two distributions with equal prior probabilities into one of the two classes,

$$\widehat{\mathcal{M}}_n(\boldsymbol{z}) = \begin{cases} 1, & \lambda(\boldsymbol{z}, \boldsymbol{\sigma}_n) > 0 \\ 2, & \lambda(\boldsymbol{z}, \boldsymbol{\sigma}_n) \leq 0 \end{cases},$$

where $\lambda(\boldsymbol{z}, \boldsymbol{\sigma}_n) = \sum_{j=1}^{p} \{\Psi_{1j}(z_j, \sigma_{1j,n}) - \Psi_{2j}(z_j, \sigma_{2j,n})\}$ and $\Psi_{rj}(z_j, \sigma_{rj,n}) = K\left(\frac{z_j - \hat{\delta}_{rj}}{\sigma_{rj,n}}\right)$, δ_{rj} is the marginal mode of F_{rj}, $\hat{\delta}_{rj}$ is obtained based on the samples. Given the observed data, the performance of the classification rule is measured by the misclassification error

$$R(\widehat{\mathcal{M}}) = P\{\text{label}(\boldsymbol{z}) \neq \hat{\mathcal{M}}(\boldsymbol{z})\} = \pi_1 \int I\{\lambda(\boldsymbol{z}, \boldsymbol{\sigma}) \leq 0\} d\mathcal{F}_1(\boldsymbol{z}) + \pi_2 \int I\{\lambda(\boldsymbol{z}, \boldsymbol{\sigma}) > 0\} d\mathcal{F}_2(\boldsymbol{z}),$$

where $\boldsymbol{z}$ is independent of the observed $\mathbf{X}$s, label($\boldsymbol{z}$) denotes the true class of $\boldsymbol{z}$. In the following proofs, we restrict ourselves to assume that $\pi_1 = \pi_2$.

S2.1 Notation and definitions

We first introduce some notations, definitions and basic facts that will be used in the paper.

For two sequences of positive numbers a_n and b_n, $a_n \lesssim b_n$ means that for some constant $c > 0$, $a_n \leq cb_n$ for all n, and $a_n \asymp b_n$ if $a_n \lesssim b_n$ and $b_n \lesssim a_n$. Denote by $\mathcal{B}(x, r)$ the closed Euclidean ball in $\mathbb{R}^p$ centered at $x \in \mathbb{R}^p$ and of radius $r > 0$. For any multi-index $s = (s_1, \ldots, s_p) \in \mathcal{N}^p$ and any $x = (x_1, \ldots, x_p) \in \mathbb{R}^p$, we denote $|s| = \sum_{i=1}^p s_i$, $s! = s_1! \cdots s_p!$, $x^s = x_1^{s_1} \cdots x_p^{s_p}$ and $\|x\| := (x_1^2 + \cdots x_p^2)^{1/2}$. Let D^s denote the differential operator $D^s := \frac{\partial^{s_1+\cdots+s_p}}{\partial x_1^{s_1} \cdots \partial x_p^{s_p}}$.

Let $\beta > 0$, denote by $\lfloor \beta \rfloor$ the maximal integer that is strictly less than β. For any $x \in \mathbb{R}^p$ and any $\lfloor \beta \rfloor$-times continuously differentiable real-valued function g on $\mathbb{R}^p$, we denote by its Taylor polynomial of degree $\lfloor \beta \rfloor$ at point x,

$$g_x(x') := \sum_{|s| \leq \lfloor \beta \rfloor} \frac{(x' - x)^s}{s!} D^s g(x).$$

Let $L > 0$. The $(\beta, L, \mathbb{R}^p)$-Hölder class of functions, denoted by $\Sigma(\beta, L, \mathbb{R}^p)$, is defined as the set of functions $g : \mathbb{R}^p \to \mathbb{R}$ that are $\lfloor \beta \rfloor$ times continuously differentiable and satisfy, for any $x, x' \in \mathbb{R}^p$, the inequality

$$|g(x') - g_x(x')| \leq L\|x - x'\|^\beta.$$

We finally recall some notations related to unimodal estimators. We use $\lambda(\boldsymbol{z}, \sigma_n)$ to estimate $\eta(\boldsymbol{z}) - 1/2$ or $f_1(\boldsymbol{z}) - f_2(\boldsymbol{z}) \geq 0$. For unimodal distributions $\mathcal{F}_{rj}$, $\sigma_{rj,n} > 0$, $\boldsymbol{z} = (z_1, \ldots, z_p)^T$, denote by $\hat{\delta}_{rj}$ which maximizes ${n_r}^{-1} \sum_{i=1}^{n_r} K_{\sigma_{rj,n}}(X_{ij}^{(r)} - \hat{\delta}_{rj})$ for $r = 1, 2$ and $j = 1, \ldots, p$. The unimodal estimator $\hat{\eta}_n(\boldsymbol{z})$ is defined by $\sum_{j=1}^p \left\{ K(\frac{z_j - \hat{\delta}_{1j}}{\sigma_{1j,n}}) - K(\frac{z_j - \hat{\delta}_{2j}}{\sigma_{2j,n}}) \right\} + 1/2$. The value $\sigma_{rj,n}$ is called the bandwidth and the function K is called the kernel function, in this paper we use Gaussian kernel.

S2.2 Convergence rates for fixed p

For any classifier $\mathcal{M}$, $R(\mathcal{M}) - R(\mathcal{M}^*) = \frac{1}{2} d(\mathcal{M}, \mathcal{M}^*)$, where $d(M_1, M_2) = \int_{M_1 \triangle M_2} I\{|f_1(x) - f_2(x)| \geq 0\} dx$ is a distance defined on measurable subsets of $\mathbb{R}^p$, $M_1 \triangle M_2 = (M_1^c \cap \mathcal{M}_2) \cup (M_1 \cap M_2^c)$ is the symmetric difference of M_1 and M_2. Since the densities of f_1 and f_2 are assumed to be unknown, the Bayesian rule $\mathcal{M}^*$ is not available, one has to use empirical rules, such as unimodal classification rules $\hat{\mathcal{M}}$. There is another natural way of measuring the accuracy of a decision rule $\mathcal{M}$ through the quantity:

$$d_\triangle(\mathcal{M}, \mathcal{M}^*) = \int_{\mathcal{M} \triangle \mathcal{M}^*} dP,$$

where P is the probability measure.

Let $\mathcal{K} = \prod_{i=1}^{p} K_j : \mathbb{R}^p \to \mathbb{R}$ be a p-dimensional function defined as the product of p unidimensional functions K_j. Denote by $h = (h_1, \ldots, h_p)$ a set of bandwidths, then we have for all $\mathcal{M}$, the risk $R(\mathcal{M})$ can be estimated by

$$R^h(\mathcal{M}) = \frac{1}{2}\left[\frac{1}{m}\sum_{i=1}^{m} G_h(\mathbf{X}_i) + \frac{1}{n}\sum_{i=1}^{n} G_h(\mathbf{Y}_i)\right],$$

where for a given $z \in \mathbb{R}^p$, $G_h(z) = \int \frac{1}{h}\mathcal{K}(\frac{z-x}{h})dx$. Define $\hat{\mathcal{M}}^h = \arg\min_{\mathcal{M}} R^h(\mathcal{M})$, where $h = (h_1, \ldots, h_d)$ has to be chosen explicitly. Denote $R_E^h(\cdot) := ER^h(\cdot)$:

$$\begin{aligned} R(\widehat{\mathcal{M}}) - R(\mathcal{M}^*) &\leq R(\widehat{\mathcal{M}}) - R^h(\mathcal{M}) + R^h(\mathcal{M}) - R(\mathcal{M}^*) \\ &\leq \sup |R_E^h - R^h|(\mathcal{M}, \mathcal{M}^*) + \sup |R_E^h - R^h|(\mathcal{M}, \mathcal{M}^*), \end{aligned}$$

where for any M, M': $|R_E^h - R^h|(M, M') = |R_E^h(M) - R_E^h(M') - R^h(M) + R^h(M')|$, and similarly, $|R_E^h - R|(M, M') = |R_E^h(M) - R_E^h(M') - R(M) + R(M')|$.

Assumption (MA). There exist constants $C_0 > 0$ and $\alpha > 0$ such that

$$P\left(0 < \Big|\sum_{j=1}^{p}\{\Psi_{1j}(z_j, \sigma_{1j,n}) - \Psi_{2j}(z_j, \sigma_{2j,n})\}\Big| < t\right) \leq C_0 t^{\alpha} \qquad \forall t > 0. \tag{S2.27}$$

The case $\alpha = 0$ is trivial and is included for notational convenience. The other extreme case $\alpha = \infty$ is most advantageous for classification: the two classes are well separated. Assumption (MA) provides a useful characterization of the decision rule $\lambda(\boldsymbol{z}, \sigma_n)$ in the vicinity of the level $\lambda(\boldsymbol{z}, \sigma_n) = 0$, which turns out to be crucial for convergence of classifiers.

Lemma S7. *Let $\hat{\eta}_n$ be an estimator of the regression function η and $\mathcal{P}$ a set of probability distributions on $\mathcal{Z}$ such that for some constants $C_1 > 0$, $C_2 > 0$, for some positive sequence a_n, for $n \geq 1$ and any $\epsilon > 0$, and for almost all x with respect to (w.r.t.) $\mathcal{F}_r$ for $r = 1, 2$, we have*

$$\sup_{P \in \mathcal{P}} P(|\hat{\eta}_n(z) - \eta(z)| \geq \epsilon) \leq C_1 \exp(-C_2 a_n \epsilon^2). \tag{S2.28}$$

Considering the classifier $\hat{\mathcal{M}}_n = I\{\hat{\eta}_n(z) \geq 1/2\}$. If all the distributions $P \in \mathcal{P}$ satisfy the margin Assumption (MA), we have

$$\sup_{P \in \mathcal{P}} \{ER(\hat{\mathcal{M}}_n) - R(\mathcal{M}^*)\} \leq C a_n^{-(1+\alpha)/2}$$

for $n \geq 1$ with some constant $C > 0$ depending only on α, C_0, C_1 and C_2.

Proof. Consider the sets $A_j \subset \mathbb{R}^p$, $j = 1, 2, \ldots$ defined as

$$A_0 := \{x \in \mathbb{R}^p : 0 < |\eta(z) - 1/2| \leq \epsilon\},$$

$$A_j := \{x \in \mathbb{R}^p : 2^{j-1}\delta < |\eta(z) - 1/2| \leq 2^j \delta\} \quad \text{for} j \geq 1.$$

For any $\epsilon > 0$, we may write

$$\begin{aligned} ER(\widehat{\mathcal{M}}_n) - R(\mathcal{M}^*) &= E(|2\eta(Z) - 1| I\{\hat{\mathcal{M}}_n(Z) \neq \mathcal{M}^*(Z)\}) \\ &= \sum_{j=0}^{\infty} E(|2\eta(Z) - 1| I\{\widehat{\mathcal{M}}_n(Z) \neq \mathcal{M}^*(Z)\} I(Z \in A_j)) \\ &\leq 2\epsilon P(0 < |\eta(Z) - 1/2| \leq \epsilon) \qquad \text{(S2.29)} \\ &\quad + \sum_{j \geq 1} E(|2\eta(Z) - 1| I\{\hat{\mathcal{M}}_n(Z) \neq \mathcal{M}^*(Z)\} I(Z \in A_j)). \end{aligned}$$

On the event $\{\hat{\mathcal{M}}_n \neq \mathcal{M}^*\}$ we have $|\eta - 1/2| \leq |\hat{\eta}_n - \eta|$. So for any $j \geq 1$, we get

$$\begin{aligned} & E(|2\eta(Z) - 1| I\{\hat{\mathcal{M}}_n(Z) \neq \mathcal{M}^*(Z)\} I(Z \in A_j)) \\ \leq \; & 2^{j+1}\epsilon E[I\{|\hat{\eta}_n(Z) - \eta(Z)| \geq 2^{j-1}\epsilon\} I\{0 < |\eta(Z) - 1/2| \leq 2^j \epsilon\}] \\ \leq \; & 2^{j+1}\epsilon E_Z[P\{|\hat{\eta}_n(Z) - \eta(Z)| \geq 2^{j-1}\epsilon\} I\{0 < |\eta(Z) - 1/2| \leq 2^j \epsilon\}] \\ \leq \; & C_1 2^{j+1}\epsilon \exp(-C_2 a_n (2^{j-1}\epsilon)^2) P_Z\{0 < |\eta(Z) - 1/2| \leq 2^j \epsilon\} \\ \leq \; & 2C_1 C_0 2^{j(1+\alpha)} \epsilon^{1+\alpha} \exp(-C_2 a_n (2^{j-1}\epsilon)^2), \end{aligned}$$

where in the last inequality we have used Assumption (MA). Now, from inequality (S2.29), taking $\epsilon = a_n^{-1/2}$ and using Assumption (MA) to bound the first term of the right-hand side of (S2.29), we get

$$\begin{aligned} ER(\widehat{\mathcal{M}}_n) - R(\mathcal{M}^*) & \leq 2C_0 a_n^{-(1+\alpha)/2} + C a_n^{-(1+\alpha)/2} \sum_{j \geq 2} 2^{j(1+\alpha)} \exp(-C_2 2^{2j-2}) \\ & \leq C a_n^{-(1+\alpha)/2}. \end{aligned}$$

□

Lemma S8. *Assume that the cumulative distribution functions of the two populations are such that the density functions $f_1(z) = \mathcal{F}_1'(z)$ and $f_2(z) = \mathcal{F}_2'(z)$ exist for z and are nonzero on the same domain. Further assume that there is a point z_0 with $\pi_1 f_1(z_0) = \pi_2 f_2(z_0)$ so that $\pi_1 f_1(z) > \pi_2 f_2(z)$ for z on one side of z_0 and $\pi_1 f_1(z) > \pi_2 f_2(z)$ for z on the other side of z_0. Then the unimodal classifier using the mode $\tilde{\delta}_{rj}$ that minimizes the true misclassification probability achieves the Bayes* *misclassification probability.*

The proof of this lemma is easy, we refer to the details of Lemma 2 of supplementary materials of Hennig and Viroli (2016). In addition, assume that K is a kernel satisfying that: (1) $\exists c > 0$: $K(x) \geq c$ if $|x| \leq c$, (2) $\int K(u)du = 1$, (3) $\int (1 + u^{4\beta})K^2(u)du < \infty$, (4) $\sup_u (1 + |u|^{2\beta})K(u) < \infty$.

Let $\sigma_{rj,n} > 0$, denote by $\hat{\mathcal{M}}_n(z) = \lambda(z, \sigma_n)$ the unimodal classifier with bandwidth $\sigma_{rj,n} > 0$ and kernel K satisfying the above conditions. Denote by $\mathcal{M}^*(z)$ the classifier achieves the optimal Bayes risk. Then we have the following theorem.

Theorem S1. *Let $\mathcal{P}$ be a class of probability distributions P on $\mathcal{Z}$ such that the regression function η belongs to the Hölder class $\Sigma(\beta, L, \mathbb{R}^p)$ and the marginal law of X satisfies the strong density assumption. Let $\sigma = \max\{\sigma_{rj,n}\}$. Then there exist constant $C_1, C_2, C_3 > 0$ such that for any $0 < \sigma \leq r_0/c$, any $C_3\sigma^\beta < \epsilon$ and any $m, n \geq 1$, the unimodal estimator $\hat{\mathcal{M}}_n$ satisfies*

$$\sup_{P \in \mathcal{P}} P(|\hat{\mathcal{M}}_n(z) - \mathcal{M}^*(z)| \geq \epsilon) \leq C_1 \exp(-C_2(m \wedge n)\sigma^p \epsilon^2) \tag{S2.30}$$

for almost all x w.r.t P_Z. As a consequence, there exist $C_1, C_2 > 0$ such that for $\sigma = n^{-1/(2\beta+p)}$ and any $\epsilon > 0$, $n \geq 1$ we have

$$\sup_{P \in \mathcal{P}} P(|\hat{\mathcal{M}}_n(z) - \mathcal{M}^*(z)| \geq \epsilon) \leq C_1 \exp(-C_2 n^{2\beta/(2\beta+p)} \epsilon^2) \tag{S2.31}$$

for almost all x w.r.t P_X. The constants C_1, C_2, C_3 depend only on β, p, L.

Proof. Consider the matrix $B := (B_{s_1,s_2})$ with elements $B_{s_1,s_2} := \int u^{s_1+s_2} hK(u)du$. Define

$$W_i := \frac{1}{h^d}(\frac{z_i - z}{h})^{s_1+s_2} K(\frac{z_i - z}{h}) - \int hu^{s_1+s_2} K(u)du$$

We have $EW_i = 0$, $|W_i| \le h^{-d} \sup_{u\in\mathbb{R}^d}(1 + \|u\|^{2\beta})K(u) := \kappa_1 h^{-d}$ and

$$Var(W_i) \le \frac{1}{h^{2d}} E(\frac{z_i - z}{h})^{2(s_1+s_2)} K^2(\frac{z_i - z}{h}) \le \frac{1}{h^d} \int (1 + \|u\|^{4\beta})K^2(u) := \frac{\kappa_2}{h^d}.$$

Using Bernstein's inequality, for any $\varepsilon > 0$, we have

$$P(|\bar{B} - B| > \varepsilon) = P\left(|\frac{1}{n} \sum_{i=1}^{n} T_i| > \varepsilon \right) \le 2 \exp\left\{ - \frac{(n)h^d \varepsilon^2}{2\kappa_2 + 2\kappa_1 \varepsilon/3} \right\}.$$

This implies that $P(\lambda_{\bar{B}} \le \mu_0) \le 2M^2 \exp(-Cnh^d)$, where M^2 is the number of elements of the matrix $\bar{B}$. Assume in what follows that n is large enough so that $\mu_0 > (\log n)^{-1}$. Then for $\lambda_{\bar{B}} > \mu_0$ we have $|\hat{\mathcal{M}}_n(z) - \mathcal{M}^*(z)| \le |\hat{\mathcal{M}}_n(z) - \hat{\mathcal{M}}^*(z)|$. Therefore,

$$\begin{aligned} P(|\hat{\mathcal{M}}_n^*(z) - \mathcal{M}(z)| \ge \epsilon) \;\; &\le \;\; P(\lambda_{\bar{B}} \le \mu_0) \\ &\quad + P(|\hat{\mathcal{M}}_n(z) - \hat{\mathcal{M}}^*(z)| \ge \epsilon, \lambda_{\bar{B}} > \mu_0). \end{aligned}$$

We now evaluate the second probability. For $\lambda_{\bar{B}} \le \mu_0$, introduce

$$T_{1i} := \frac{1}{mh^p} \sum_{i=1}^{m} \mathcal{K}\left(\frac{X_i - z}{h} \right) - \sum_{j=1}^{p} \frac{1}{\sigma_j} K\left(\frac{z_j - \delta_{1j}(X_j)}{\sigma_j} \right),$$

$$T_{2i} := \frac{1}{nh^p} \sum_{i=1}^{n} \mathcal{K}\left(\frac{Y_i - z}{h} \right) - \sum_{j=1}^{p} \frac{1}{\sigma_j} K\left(\frac{z_j - \delta_{2j}(Y_j)}{\sigma_j} \right),$$

then $|\hat{\mathcal{M}}_n(z) - \hat{\mathcal{M}}^*(z)| \leq |T_{1i}| + |T_{2i}|$. Rewrite T_{1i} as,

$$T_{1i} = \frac{1}{m}\sum_{i=1}^{m}\left\{\frac{1}{h^p}\mathcal{K}\left(\frac{X_i - z}{h}\right) - \sum_{j=1}^{p}\frac{1}{\sigma_j}K\left(\frac{z_j - \delta_{1j}(X_j)}{\sigma_j}\right)\right\} := \frac{1}{m}\sum_{i=1}^{m} a_{1i},$$

similarly,

$$T_{2i} = \frac{1}{n}\sum_{i=1}^{n}\left\{\frac{1}{h^p}\mathcal{K}\left(\frac{Y_i - z}{h}\right) - \sum_{j=1}^{p}\frac{1}{\sigma_j}K\left(\frac{z_j - \delta_{1j}(Y_j)}{\sigma_j}\right)\right\} := \frac{1}{n}\sum_{i=1}^{n} a_{2i},$$

We have $Ea_{ri} = 0$, $|a_{ri}| \leq h^{-p}\sup_{u\in\mathbb{R}^d}(1 + \|u\|^{2\beta})\mathcal{K}(u) := \kappa_1 h^{-p}$ and the following bound for the variance of T_i:

$$Var(a_{ri}) \leq \frac{1}{h^{2p}}E\mathcal{K}^2\left(\frac{X - z}{h}\right) = \frac{1}{h^p}\int \mathcal{K}^2(u)du \leq \kappa_2 h^{-p}.$$

Using Bernstein's inequality, for any $\epsilon_1, \epsilon_2 > 0$, we obtain

$$P(|T_{1i}| \geq \epsilon_1) \leq 2\exp\left\{-\frac{mh^p\epsilon_1^2}{\kappa_2/2 + 2\kappa_1\epsilon_1/3}\right\}, \quad P(|T_{2i}| \geq \epsilon_2) \leq 2\exp\left\{-\frac{nh^p\epsilon_2^2}{\kappa_2/2 + 2\kappa_1\epsilon_2/3}\right\}.$$

It follows that

$$P(|\hat{\mathcal{M}}_n^*(z) - \mathcal{M}(z)| \geq \epsilon) \leq C_1\exp(-C_2(m \wedge n)h^p\epsilon^2).$$

□

By using Theorem S1, it follows directly that the convergence rate for unimodal classifier when p is fixed.

Theorem S2. *(Convergence rate when p is fixed) For any $m, n \geq 1$, the excess risk of the unimodal classifier $\hat{\mathcal{M}}_n$ with bandwidth $\sigma = n^{-1/(2\beta+p)}$ satisfies*

$$\sup\{ER(\hat{\mathcal{M}}_n) - R(\mathcal{M}^*)\} \leq Cn^{-\beta(1+\alpha)/(2\beta+p)},$$

where the constant $C > 0$ depends only on α, C_0, C_1 and C_2.

S2.3 Convergence rate when $p \to \infty$

In fact, the convergence rate of the mode-based classifiers depend on the assumptions on the joint distributions of $(\mathbf{X}, \mathbf{Y}, \mathcal{C})$. A standard way to derive the convergence rate is to introduce a class of joint distributions of $(\mathbf{X}, \mathbf{Y}, \mathcal{C})$ and to declare $\tilde{\mathcal{M}}_n$ optimal if it achieves the best rate of convergence in a minimax sense on this class. Since the distribution of $(\mathbf{X}, \mathbf{Y})$ is unknown, it is rather difficult to obtain the convergence rate. We consider a specific normal case in the following, that is, $\mathbf{X}$ and $\mathbf{Y}$ both come from normal distributions.

Suppose one observes n_1 samples from class $1 : X_1, \ldots, X_m \overset{i.i.d}{\sim} \mathcal{F}_1(\boldsymbol{\delta}_1)$, n_2 samples from class $2 : Y_1, \ldots, Y_n \overset{i.i.d}{\sim} \mathcal{F}_2(\boldsymbol{\delta}_2)$, $\boldsymbol{\delta}_1, \boldsymbol{\delta}_2$ are the mode vector of $\mathcal{F}_1$ and $\mathcal{F}_2$, and the prior probabilities of class 1 and class 2 are π_1 and π_2, respectively. The goal is to construct a classification rule $\hat{\mathcal{M}}$, which is a function of X_i's and Y_i's, to classify a future data point $\boldsymbol{z} \sim \pi_1 \mathcal{F}_1(\boldsymbol{\delta}_1) + \pi_2 \mathcal{F}_2(\boldsymbol{\delta}_2)$. The parameters of this model is collected by $\Theta = (\pi_1, \pi_2, \boldsymbol{\delta}_1, \boldsymbol{\delta}_2)$. Let $n = \min\{n_1, n_2\}$. For any classification rule $\hat{\mathcal{M}} : \mathbb{R}^p \to \{1, 2\}$, the accuracy is measured by the classification error

$$R_\Theta(\hat{\mathcal{M}}) = E_\Theta[I\{\hat{\mathcal{M}}(\boldsymbol{z}) \neq L(\boldsymbol{z})\}], \tag{S2.32}$$

where $L(\boldsymbol{z})$ denotes the true class label of $\boldsymbol{z}$, that is, $L(\boldsymbol{z}) = 1$ if $\boldsymbol{z} \sim \mathcal{F}_1(\boldsymbol{\delta}_1)$ and 2 otherwise. In the ideal setting where all the parameters Θ and $\mathcal{F}_1$, $\mathcal{F}_2$ are known, such as $\mathcal{F}_1 = \mathcal{N}_p(\boldsymbol{\delta}_1, I_p)$ and $\mathcal{F}_2 = \mathcal{N}_p(\boldsymbol{\delta}_2, I_p)$, by using a Gaussian kernel function (L_2 distance metric), the optimal classifier in this case is reduced to the linear discriminant rule (LDA) is

$$\mathcal{M}_\theta^*(\boldsymbol{z}) = \begin{cases} 1, & -\Delta^T(\boldsymbol{z} - \bar{\boldsymbol{\delta}}) + \log(\frac{\pi_1}{\pi_2}) \leq 0, \\ 2, & -\Delta^T(\boldsymbol{z} - \bar{\boldsymbol{\delta}}) + \log(\frac{\pi_1}{\pi_2}) > 0, \end{cases} \tag{S2.33}$$

where $\Delta = \boldsymbol{\delta}_2 - \boldsymbol{\delta}_1$, $\bar{\boldsymbol{\delta}} = \frac{1}{2}(\boldsymbol{\delta}_1 + \boldsymbol{\delta}_2)$. The oracle classification rule in (S2.33) is the Bayes rule and achieves the minimal classification error. For ease of presentation, let us define the discriminant function by

$$\lambda(\boldsymbol{z}; \Theta) = -2\Delta^T(\boldsymbol{z} - \bar{\boldsymbol{\delta}}) + 2\log(\pi_1/\pi_2). \tag{S2.34}$$

Then $\lambda(\boldsymbol{z}; \Theta) = 0$ characterizes the classification boundary of the oracle componentwise unimodal based rule, and (S2.33) can be written as

$$\mathcal{M}_\Theta^*(\boldsymbol{z}) = 1 + I\{\lambda(\boldsymbol{z}; \Theta) \leq 0\},$$

and $R_\Theta(\mathcal{M}_\Theta^*) = \min_{\mathcal{M} \in M} R_\Theta(\mathcal{M})$, where M is the set of all classification rules.

First, we consider $\pi_1 = \pi_2 = 1/2$, let $\alpha = \sqrt{\Delta^T \Delta}$, then consider a collection of the parameter spaces $\mathcal{G}(s, D_{n,p})$ defined by

$$\mathcal{G}(s, D_{n,p}) = \{\Theta = (1/2, 1/2, \boldsymbol{\delta}_1, \boldsymbol{\delta}_2) : \boldsymbol{\delta}_1, \boldsymbol{\delta}_2 \in \mathbb{R}^p, \|\Delta\|_0 \leq s, D_{n,p} \leq \alpha \leq 3D_{n,p}\},$$

where $D_{n,p} > 0$ can potentially grow with n and p. The sparsity constraint $\|\Delta\|_0 \leq s$ implies that only a limited number of covariates have discriminating power and contribute to the classification task, which is reasonable in high-dimensional data analysis. In this case, $R_{opt}(\Theta) = \Phi(-\frac{1}{2}\alpha)$, where Φ is the cumulative distribution function of the standard normal distribution.

Theorem S3. *Consider the parameter space $\mathcal{G}(s, D_{n,p})$, s and p approach ∞ as $n \to \infty$, and*

$$D_{n,p} = o\left[\sqrt{\frac{n}{s\log(p)}}\right]$$

with $n \to \infty$.

(a) If $D_{n,p}$ is a fixed constant not depending on n and p, then, for any constant $\gamma \in (0,1)$, we have

$$\inf\left\{\gamma : \inf_{\hat{\mathcal{M}}} \sup_{\Theta\in\mathcal{G}(s,D_{n,p})} P\{R_\Theta(\hat{\mathcal{M}}) - R_\Theta(\mathcal{M}^*_\Theta)\} \le 1-\gamma\right\} \asymp \frac{s\log(p)}{n}.$$

(b) If $D_{n,p} \to \infty$ as $n \to \infty$, then for sufficiently large n and any constant $\beta \in (0,1)$,

$$\inf\left\{\gamma : \inf_{\hat{\mathcal{M}}} \sup_{\Theta\in\mathcal{G}(s,D_{n,p})} P\{R_\Theta(\hat{\mathcal{M}}) - R_\Theta(\mathcal{M}^*_\Theta)\} \le 1-\gamma\right\} \asymp \frac{s\log(p)}{n} \exp\left[-\left\{\frac{1}{8}+o(1)\right\}D^2_{n,p}\right].$$

It is worth noting that $D_{n,p}$ represents the magnitude of α, which is interpreted as the signal-to-noise ratio. As shown in the second case, when the signal-to-noise ratio grows, the classification problem becomes easier.

Lemma S9. *Suppose that $\mathbf{X}_1, \ldots, \mathbf{X}_n \overset{iid}{\sim} N_p(\boldsymbol{\delta}, I_p)$, and assume that $\hat{\boldsymbol{\delta}}$ is the sample mode. Let $\Gamma(s) = \{\boldsymbol{u} \in \mathbb{R}^p : \|\boldsymbol{u}_{S^c}\|_1 \le \|\boldsymbol{u}_S\|_1,$ for some $S \subset [p]$ with $|S| = s\}$; then,* *with probability at least* $1 - p^{-1}$,

$$\sup_{\boldsymbol{u} \in \Gamma(s)} \boldsymbol{u}^T (\hat{\boldsymbol{\delta}} - \boldsymbol{\delta}) \lesssim \sqrt{\frac{s \log(p)}{n}}.$$

Lemma S10. *Suppose that* $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^p$. *Let* $h = \boldsymbol{x} - \boldsymbol{y}$ *and* $\mathcal{S} = supp(\boldsymbol{y})$. *If* $\|\boldsymbol{x}\|_1 \leq \|\boldsymbol{y}\|_1$, *then* $h \in \Gamma(s)$ *with* $s = |S|$, *i.e.*

$$\|h_{\mathcal{S}^c}\|_1 \leq \|h_{\mathcal{S}}\|_1.$$

For a vector $\mathbf{x} \in \mathbb{R}^p$, define $\|\mathbf{x}\|_{2,s} = \sup_{\|\boldsymbol{y}\|_2 = 1, \boldsymbol{y} \in \Gamma(s)} |\mathbf{x}^T \boldsymbol{y}|$

Proof. Proof of Theorem S3. Let $\kappa_n = \|\hat{\boldsymbol{\delta}}_1 - \boldsymbol{\delta}_1\|_{2,s} \vee \|\hat{\boldsymbol{\delta}}_2 - \boldsymbol{\delta}_2\|_{2,s}$. We shall show that

$$R_\theta(\hat{\mathcal{M}}) - R_{opt}(\Theta) \lesssim \exp(-\alpha^2/8)\alpha\kappa_n^2.$$

Given the estimator $\boldsymbol{\delta}_k$, the sample $\boldsymbol{z}$ is classified as

$$\hat{\mathcal{M}}(\boldsymbol{z}) = \begin{cases} 1, & (\boldsymbol{z} - (\hat{\boldsymbol{\delta}}_1 + \hat{\boldsymbol{\delta}}_2)/2)^T \hat{\Delta}^T \geq 0, \\ 2, & (\boldsymbol{z} - (\hat{\boldsymbol{\delta}}_1 + \hat{\boldsymbol{\delta}}_2)/2)^T \hat{\Delta}^T < 0. \end{cases}$$

Let $\hat{\alpha} = \sqrt{\hat{\boldsymbol{\delta}}^T \hat{\boldsymbol{\delta}}}$ and $\hat{\boldsymbol{\delta}} = (\hat{\boldsymbol{\delta}}_1 + \hat{\boldsymbol{\delta}}_2)/2$. The misclassification error is

$$R_\Theta(\hat{C}) = \frac{1}{2}\Phi\left(- \frac{(\hat{\boldsymbol{\delta}} - \boldsymbol{\delta}_1)^T \hat{\Delta}}{\hat{\alpha}} \right) + \frac{1}{2}\bar{\Phi}\left(- \frac{(\hat{\boldsymbol{\delta}} - \boldsymbol{\delta}_1)^T \hat{\Delta}}{\hat{\alpha}} \right),$$

with $R_{opt}(\Theta) = \frac{1}{2}\Phi(-\alpha/2) + \frac{1}{2}\bar{\Phi}(\alpha/2)$. Define an intermediate quantity

$$R^* = \frac{1}{2}\Phi\left(- \frac{\Delta^T \hat{\Delta}/2}{\hat{\alpha}} \right) + \frac{1}{2}\bar{\Phi}\left(\frac{\Delta^T \hat{\Delta}/2}{\hat{\alpha}} \right)$$

We first show that $R^* - R_{opt}(\Theta) \lesssim \exp(-\alpha^2/8)\alpha\kappa_n^2$. Applying Taylor's expansion to the two terms in R^* at $-\alpha/2$ and $\alpha/2$, we have

$$R^* - R_{opt}(\Theta) = \frac{1}{2}\left(\frac{\alpha}{2} - \frac{\Delta^T\hat{\Delta}}{2\hat{\alpha}}\right)\Phi'\left(\frac{\alpha}{2}\right) + \frac{1}{2}\left(\frac{\alpha}{2} - \frac{\Delta^T\hat{\Delta}}{2\hat{\alpha}}\right)\Phi'\left(-\frac{\alpha}{2}\right) + O\left\{\exp(-\frac{\alpha^2}{8})\frac{1}{\alpha}\kappa_n^4\right\} \tag{S2.35}$$

The remaining term can be written as

$$\frac{1}{2}\left(\frac{\alpha}{2} - \frac{\Delta^T\hat{\Delta}}{2\hat{\alpha}}\right)^2\Phi''(t_{1,n}) + \frac{1}{2}\left(\frac{\alpha}{2} - \frac{\Delta^T\hat{\Delta}}{2\hat{\alpha}}\right)^2\Phi''(t_{2,n}),$$

where $t_{1,n}$ and $t_{2,n}$ are some constants satisfying that $|t_{1,n}|$ and $|t_{2,n}|$ are between $\Delta/2$ and $\Delta^T\hat{\Delta}/(2\hat{\alpha})$. Therefore, the remaining term can be bounded by using the facts that

$$\left|\frac{\Delta^T\hat{\Delta}}{2\hat{\alpha}} - \frac{\alpha}{2}\right| = O\left(\frac{1}{\alpha}\kappa_n^2\right)$$

and

$$\Phi''(t_n) = O\{\exp(-\frac{\alpha^2}{8})\alpha\},$$

for $|t_n|$ is between $\alpha/2$ and $\Delta^T\hat{\Delta}/(2\hat{\alpha})$. In fact, for the first term, we have

$$\left|\alpha - \frac{\Delta^T\hat{\Delta}}{\hat{\alpha}}\right| = \left|\|\Delta\|_2 - \frac{\Delta^T\hat{\Delta}}{\|\hat{\Delta}\|_2}\right| = \left|\frac{\|\Delta\|_2\|\hat{\Delta}\|_2 - \Delta^T\hat{\Delta}}{\|\hat{\Delta}\|_2}\right| \lesssim \frac{1}{\alpha}\|\hat{\Delta} - \Delta\|_2^2 \lesssim \frac{1}{\alpha}\kappa_n^2.$$

In addition, $|\Phi''(t_n)| \asymp \alpha\exp\{-(\alpha/2)^2/2\} = \alpha\exp(-\alpha^2/8)$. □

Then equation (S2.35) can be further expanded such that

$$\begin{aligned}
R^* - R_{opt}(\Theta) &\asymp \frac{1}{2}\bigg(\frac{\alpha}{2} - \frac{\Delta^T\hat{\Delta}}{2\hat{\alpha}}\bigg)\exp\bigg(-\frac{1}{2}(\frac{\alpha}{2})^2\bigg) + \frac{1}{2}\bigg(\frac{\alpha}{2} - \frac{\Delta^T\hat{\Delta}}{2\hat{\alpha}}\bigg)\exp\bigg(-\frac{1}{2}(-\frac{\alpha}{2})^2\bigg) \\
&\quad + O\bigg\{\exp(-\frac{\alpha^2}{8})\frac{1}{\alpha}\kappa_n^4\bigg\} \\
&= \exp(-\frac{\alpha^2}{8})(\alpha - \frac{\Delta^T\hat{\Delta}}{\hat{\alpha}}) + O\big\{\exp(-\frac{\alpha^2}{8})\frac{1}{\alpha}\kappa_n^2\big\} \\
&\lesssim \exp(-\frac{\alpha^2}{8})|\alpha - \frac{\Delta^T\hat{\Delta}}{\hat{\alpha}}| + O\big\{\exp(-\frac{\alpha^2}{8})\frac{1}{\alpha}\kappa_n^2\big\} \lesssim \exp(-\frac{\alpha^2}{8})\kappa_n^2.
\end{aligned}$$

Thus, $R^* - R_{opt}(\Theta) \lesssim \exp(-\alpha^2/8)\alpha\kappa_n^2$. To upper bound $R_\Theta(\hat{\mathcal{M}}) - R^*$, applying Taylor's expansion to $R_\Theta(\hat{\mathcal{M}})$,

$$\begin{aligned}
R_\Theta(\hat{\mathcal{M}}) = \frac{1}{2}\bigg[&\Phi(\frac{\Delta^T\hat{\Delta}/2}{\hat{\alpha}}) + \frac{(\hat{\boldsymbol{\delta}} - \boldsymbol{\delta}_1)^T\hat{\Delta} - \Delta^T\hat{\Delta}/2}{\hat{\alpha}}\Phi'(\frac{\Delta^T\hat{\Delta}/2}{\hat{\alpha}}) + O\{\exp(-\frac{\alpha^2}{8})\alpha\kappa_n^2\}\bigg] \\
-\frac{1}{2}\bigg[&\Phi(-\frac{\Delta^T\hat{\Delta}/2}{\hat{\alpha}}) + \frac{(\hat{\boldsymbol{\delta}} - \boldsymbol{\delta}_2)^T\hat{\Delta} + \Delta^T\hat{\Delta}/2}{\hat{\alpha}}\Phi'(-\frac{\Delta^T\hat{\Delta}/2}{\hat{\alpha}}) + O\{\exp(-\frac{\alpha^2}{8})\alpha\kappa_n^2\}\bigg],
\end{aligned}$$

where the remaining term can be obtained similarly. This leads to

$$\begin{aligned}
|R_\Theta(\hat{\mathcal{M}}) - R^*| &\lesssim \bigg|\bigg[\frac{\Delta^T\hat{\Delta}/2 - (\hat{\boldsymbol{\delta}} - \boldsymbol{\delta}_1)^T\hat{\Delta}}{\hat{\alpha}}\Phi'\bigg(\frac{\Delta^T\hat{\Delta}/2}{\hat{\alpha}}\bigg) \\
&\quad + \frac{\Delta^T\hat{\Delta}/2 - (\hat{\boldsymbol{\delta}} - \boldsymbol{\delta}_2)^T\hat{\Delta}}{\hat{\alpha}}\Phi'\bigg(-\frac{\Delta^T\hat{\Delta}/2}{\hat{\alpha}}\bigg) + O\{\exp(-\frac{\alpha^2}{8})\alpha\kappa_n^2\}\bigg]\bigg| \\
&= \bigg|\bigg[\frac{\Delta^T\hat{\Delta}/2 - (\hat{\boldsymbol{\delta}} - \boldsymbol{\delta}_1)^T\hat{\Delta}}{\hat{\alpha}}\exp\bigg(-\frac{1}{2}(\frac{\Delta^T\hat{\Delta}/2}{\hat{\alpha}})^2\bigg) \\
&\quad + \frac{\Delta^T\hat{\Delta}/2 - (\hat{\boldsymbol{\delta}} - \boldsymbol{\delta}_2)^T\hat{\Delta}}{\hat{\alpha}}\exp\bigg(-\frac{1}{2}(\frac{\Delta^T\hat{\Delta}/2}{\hat{\alpha}})^2\bigg) + O\{\exp(-\frac{\alpha^2}{8})\alpha\kappa_n^2\}\bigg]\bigg|.
\end{aligned}$$

Since $\Delta/2 - (\hat{\boldsymbol{\delta}} - \boldsymbol{\delta}_1) + \Delta/2 + (\hat{\boldsymbol{\delta}} - \boldsymbol{\delta}_2) = 0$, then it follows that

$$|R_\Theta(\hat{\mathcal{M}}) - R^*| \lesssim \exp\left(- \frac{\alpha^2}{8}\right)\alpha\kappa_n^2.$$

Combining all the results, we obtain

$$R_\Theta(\hat{\mathcal{M}}) - R_{opt}(\Theta) \lesssim \exp\left(- \frac{\alpha^2}{8}\right)\alpha\kappa_n^2.$$

By lemma S9, and $\kappa_n \lesssim D_{n,p}\sqrt{s\log(p)/n}$ with probability at least $1 - 12p^{-1}$, since $\alpha \in [D_{n,p}, 3D_{n,p}]$, we then have with probability at least $1 - 12p^{-1}$,

$$R_\Theta(\hat{\mathcal{M}}) - R_{opt}(\Theta) \lesssim \exp\left(- \frac{D_{n,p}^2}{8}\right)D_{n,p}^3\frac{s\log(p)}{n}.$$

(a) When $D_{n,p}$ is bounded by constant C_b, we have

$$R_\Theta(\hat{\mathcal{M}}) - R_{opt}(\Theta) \lesssim \exp\left(- \frac{D_{n,p}^2}{8}\right)\frac{s\log(p)}{n}.$$

(b) When $D_{n,p} \to \infty$ as n grows,

$$R_\Theta(\hat{\mathcal{M}}) - R_{opt}(\Theta) \lesssim \exp\left(- (\frac{1}{8} - \frac{3\log(D_{n,p})}{D_{n,p}^2})D_{n,p}^2\right)D_{n,p}^3\frac{s\log(p)}{n},$$

where $3\log(D_{n,p})/D_{n,p}^2$ is an $o(1)$ term as $n \to \infty$. This completes the proof of Theorem S3.

S3 Settings of Figure 1 and Table 1

Examples of mode-based classifier versus centroid, median and quantile-based classifiers. Each p vectors are generated from two different populations. Panel (a):

$X_j \sim 0.8N(0,1) + 0.2N(5,10^2)$, $Y_j \sim N(1,1)$; (b) $X_j \sim \chi^2(4)$, $Y_j \sim \exp(9/40)$; (c) $X_j \sim 0.3N(-10,1) + 0.4N(0,1) + 0.3N(10,1)$, $Y_j \sim X_j + 0.5$, for $j = 1, \ldots, p$. For each setting we simulate 100 observations from each class. The classification error rates and standard errors by four different classifiers with 100 replications are listed in Table 1. For each simulated dataset, we apply componentwise centroid classifier (Centroid), componentwise median classifier (Median), quantile-based classifier (quantile) and mode-based classifier (proposed) to train classifiers, we use additional 200 testing samples generated from the oracle model of each case to estimate the classification error rate. We consider $p = 10$ and 100, respectively.

S4 An illustrating example to demonstrate the selection of σ_{11}

Take the univariate populations as an example. When $p = 1$, $\sigma_{21} = \theta\sigma_{11}$ by definition, the probability of the correct classification becomes

$$\Gamma(\sigma_{11}, \theta) = \pi_1 \int I\Big\{K\Big(\frac{z-\delta_1}{\sigma_{11}}\Big) > K\Big(\frac{z-\delta_2}{\theta\sigma_{11}}\Big)\Big\} d\mathcal{F}_1(z) + \pi_2 \int I\Big\{K\Big(\frac{z-\delta_1}{\sigma_{11}}\Big) \leq K\Big(\frac{z-\delta_2}{\theta\sigma_{11}}\Big)\Big\} d\mathcal{F}_2(z). \tag{S4.36}$$

$\Gamma(\sigma_{11}, \theta)$ in (S4.36) is dominated by two parameters σ_{11} and θ. It remains the problem of how to find the two appropriate parameters to maximize the correct classification rate. We restrict ourselves to first select σ_{11} and then choose θ. Once the "best" $\hat{\sigma}_{11}$

has been determined, the optimal θ can be subsequently identified by maximizing $\Gamma(\theta, \hat{\sigma}_{11})$.

To deal with σ_{11}, we explore five different automatic and data-driven bandwidth selectors: the least squares cross validation selector (LSCV) Bowman (1984); the Sheather-Jones plug-in selector (PI) Sheather and Jones (1991); the smoothed cross validation selector (SCV), Hall et al. (1992); the normal scale bandwidth selector (NS) Chacón et al. (2011); and the Silverman's rule of thumb selector $\sigma_{\text{ROT}} = 1.06 \min\{s, \text{IQR}/1.34\} n^{-1/5}$, where s is the sample standard deviation and IQR is the corresponding interquartile range Silverman (1986). We advocate approximating the "best" σ_{11} by comparing these five different bandwidth selectors. Figure S1 shows a number of univariate examples of how the bandwidth selectors affect the misclassification rates, assuming $\pi_1 = \pi_2 = 0.5$. Four scenarios are considered here: (a) location-shifted t distributions with three degrees of freedom: t_3 and $t_3 + 1$; (b) location-shifted chi-squared distributions: χ^2_3 and $\chi^2_3 + 1$; (c) location-shifted exponential distributions: $\exp(3)$ and $\exp(3) + 1$; (d) a Gaussian distribution and a chi-squared distribution: $N(5, 1)$ and χ^2_3. For each case, we approximate σ_{11} using the above-mentioned five bandwidth selectors within training samples and then identify θ based on the misclassification rate of the determined σ_{11}. In each setting we respectively simulate $n = 40,000$ and $n = 200$ datasets to have a general idea of the effects of sample sizes on bandwidth estimation and classification accuracy to provide robust predictions. Each dataset consists of $n/2$ samples in each class, and is equally partitioned into a training set and a test set. Hence, the training set contains $n/4$ samples from $\mathcal{C}_1$ and $n/4$ samples from $\mathcal{C}_2$. We replicate this partition 40 times to obtain the mean misclassification error rate, which is presented in Figure S1. Figure S1 reveals that the five bandwidth selectors work similarly well in these four scenarios, especially for small sample size. The ROT and NS selectors are quite robust and always perform the best, whilst the LSCV breaks down if the sample size is big. Accordingly, we recommend using ROT and NS selectors to determine the best bandwidth for unimodal classifiers.

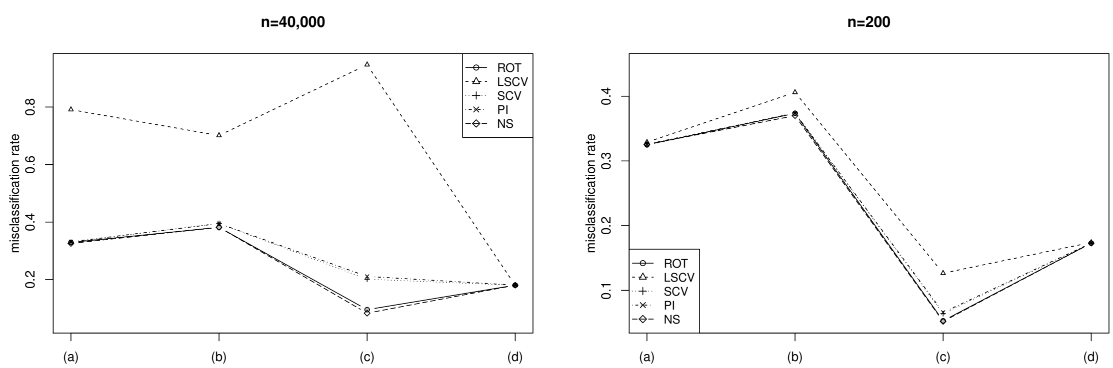

Figure S1: Misclassification rates via different bandwidth selectors for 4 scenarios with two different sample sizes. The left panel gives the results of $n = 40,000$ and the right is for $n = 200$.

S5 Detailed results of the simulation study

S5.1 Software information

In our implementation, we use the R package quantileDA for quantile, median and centroid classifiers, the package EQC for ensemble quantile classifiers, the package e1071 for the naive Bayes classifier and SVM, the package MASS for Fisher's LDA, the package class for 1-NN, the package adabag for adaptive boosting, the package xgboost for extreme gradient boosting, the package gbm for GBDT. We mainly use the default settings of these classifiers for computational and comparative purposes. For the support vector machine, tuning parameters $\gamma = (0.001, 0.01, 0.1, 1, 2)$ and Cost=(1,2,4,8,16) are optimized in each dataset according to the function tune.svm. The naiveBayes classifier has been fitted by package klaR with kernel estimate densities in stead of Gaussian densities in order to gain flexibility.

S5.2 Classification results

The tables S1–S24 show the average misclassification rates multiplied by 100 with standard errors (SD) listed in brackets, as well as the median of misclassification rates with robust standard errors (RSD) reported in brackets across different (n, p) in each setting for all three examples considered.

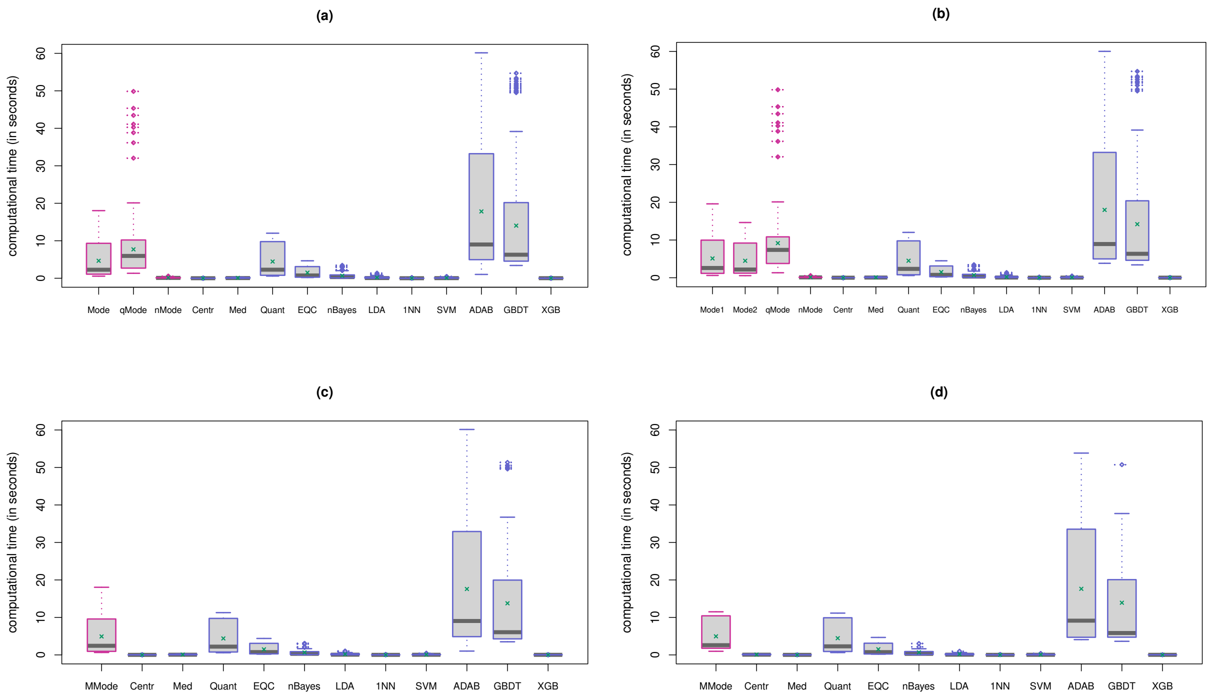

Figure S2: Computing time of the classifiers for Examples 1-3. The labels denote the different classifiers, Mode1 is simplified for unimodal classifier with ROT bandwidth selector, Mode2 for unimodal classifier with NS selector, qMode for quantile-mode classifier, nMode for naive unimodal classifier, MMode for multimodal classifier, Centr for Centroid classifier, Med for Median classifier, ADAB for AdaBoost and XGB for XGBoost. Each panel shows the distribution of the mean computational time for (a) all settings across three examples; (b) all settings of Example 1; (c) all settings of Example 2; (d) all settings of Example 3, with the cross indicating the mean.

S6 Additional Results for Real Data Analysis

For the prostate cancer dataset, to make sure of the number of modal groups for each predictor, we conduct mode testing for $p = 6,033$ genes via three testing procedures. Result are provided in Table S25 and Figure S3, respectively.

For the second dataset, i.e., the multiple myeloma dataset, Figure S4 present the KDE plots of four randomly selected genes, which is suggested to have unimodal or

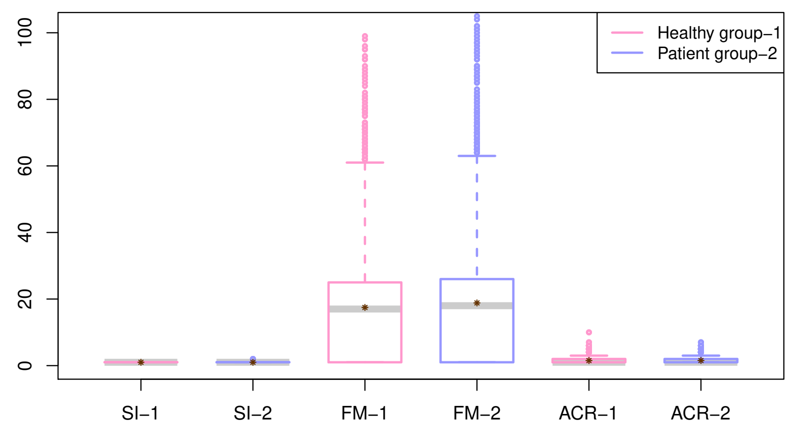

Figure S3: Box plots of the number of mode for 6033 genes determined by three mode testing procedures for the first dataset, with the star indicating the mean. The axis label "-1" denotes the healthy group and "-2" denotes the patient group.

bimodal distributions. To accurately decide on the number of modes of each predictor, we follow the same analysis path as in the first dataset and employ three mode testing procedures. Testing results are summarized in Table S26, from which the evidence of $j = 1$ mode for most of the predictors is strongly suggested. We subsequently apply our componentwise mode-based classifiers to the case where the number of modes is ascertained by each testing tool. 10-fold cross-validation is still used to evaluate the performance of the classifiers. Table S27 reports misclassification rates of all the classifiers. We find the unimodal classifiers are substantially better than other methods, and a significant performance improvement over bimodal

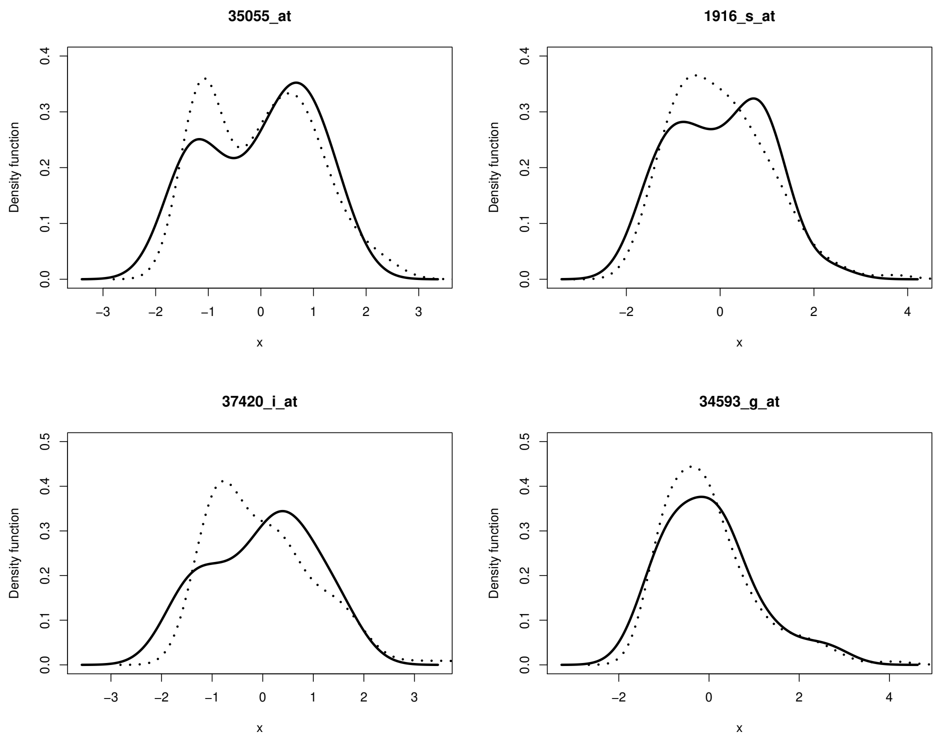

Figure S4: KDE plots of randomly selected four genes from each group. In each panel, the solid line represents the no-focal-lesion group, while the dashed line represents with-focal-lesion group.

classifiers can be made by using testing procedures, especially when the number of predictors is large. This suggests that the unimodal classifier is more applicable.

Bibliography

Bowman A.W. (1984). An alternative method of cross-validation for the smoothing of density estimates. *Biometrika*, **71**(2): 353–360.

Chacón J., Duong T., and Wand M. (2011). Asymptotics for general multivariate kernel density derivative estimators. *Statistica Sinica*, **21**(2): 807–840.

Hall P., Marron J., and Park B.U. (1992). Smoothed cross-validation. *Probability and Related Fields*, **92**(1): 1–20.

Sheather S.J. and Jones M.C. (1991). A reliable data-based bandwidth selection method for kernel density estimation. *Journal of the Royal Statistical Society: Series B (Methodological)*, **53**(3): 683–690.

Silverman B.W.(1986). *Density estimation for statistics and data analysis*. Chapman& Hall, London.

Romano J. P. (1988). *On weak convergence and Optimality of kernel density estimates of the mode*. *The Annals of Statistics*, **16**(2): 629–647.

Table S1: Misclassification rates multiplied by 100 for Case 1 of Example 1 with balanced setting

	$p = 50$			$p = 200$			$p = 1000$		
	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time
$n = 50$ (balanced)									
Unimodal(ROT)	6.15(2.06)	6.00(1.49)	0.61	0.20(0.41)	0.00(0.00)	2.29	0.00(0.00)	0.00(0.00)	12.83
Unimodal(NS)	7.18(2.01)	7.50(2.24)	0.56	0.03(0.16)	0.00(0.00)	1.73	0.00(0.00)	0.00(0.00)	10.26
Quantile-mode	2.78(1.27)	3.00(0.93)	7.74	0.00(0.00)	0.00(0.00)	7.98	0.00(0.00)	0.00(0.00)	9.58
Naive-mode	15.43(4.96)	15.00(4.66)	0.06	3.53(2.42)	3.00(2.24)	0.07	0.20(0.41)	0.00(0.00)	0.13
Centroid	20.43(3.47)	19.50(3.73)	0.01	7.40(3.44)	7.00(3.36)	0.01	16.78(18.75)	5.00(27.61)	0.02
Median	3.28(1.15)	3.00(1.49)	0.01	0.00(0.00)	0.00(0.00)	0.04	0.00(0.00)	0.00(0.00)	0.21
Quantile	5.28(2.32)	5.00(2.24)	0.58	8.40(7.12)	8.00(5.04)	2.07	24.58(10.00)	28.00(11.94)	9.63
EQC	3.10(0.63)	3.00(0.00)	0.27	0.00(0.00)	0.00(0.00)	0.64	0.00(0.00)	0.00(0.00)	2.95
nBayes	35.90(4.33)	37.00(4.48)	0.05	32.68(4.14)	32.00(3.36)	0.13	32.20(4.49)	33.00(5.22)	0.67
LDA	24.40(4.63)	24.00(3.92)	0.02	16.60(3.93)	16.00(3.92)	0.03	16.45(3.36)	16.00(3.17)	0.28
1-NN	28.15(3.94)	29.00(4.48)	0.01	20.08(5.69)	20.00(6.16)	0.01	17.40(3.69)	17.00(3.17)	0.02
SVM	8.43(2.36)	8.00(2.24)	0.01	0.00(0.00)	0.00(0.00)	0.02	0.00(0.00)	0.00(0.00)	0.13
AdaBoost	10.50(3.15)	10.50(2.99)	4.15	3.30(1.62)	3.00(1.49)	7.75	1.80(1.18)	2.00(1.49)	33.23
GBDT	6.65(2.33)	6.00(2.24)	3.86	0.65(0.70)	1.00(0.75)	4.12	0.23(0.58)	0.00(0.00)	18.22
XGBoost	32.23(3.43)	32.00(2.99)	0.01	31.10(4.98)	31.50(3.73)	0.02	29.95(4.76)	30.50(61.60)	0.02
$n = 100$ (balanced)									
Unimodal(ROT)	10.05(1.16)	10.13(1.12)	0.85	0.61(0.30)	0.50(0.19)	3.2	0.00(0.00)	0.00(0.00)	13.40
Unimodal(NS)	6.77(1.30)	6.63(1.54)	0.76	0.01(0.06)	0.00(0.00)	2.40	0.00(0.00)	0.00(0.00)	10.33
Quantile-mode	7.88(0.79)	8.00(0.93)	7.76	0.03(0.08)	0.00(0.00)	8.40	0.00(0.00)	0.00(0.00)	10.16
Naive-mode	7.25(1.31)	7.13(1.35)	0.06	0.25(0.54)	0.00(0.00)	0.09	0.00(0.00)	0.00(0.00)	0.15
Centroid	16.61(1.65)	17.00(2.24)	0.01	3.90(0.67)	3.75(0.75)	0.01	1.61(0.60)	1.50(0.56)	0.02
Median	3.77(0.72)	3.75(0.75)	0.02	0.00(0.00)	0.00(0.00)	0.05	0.00(0.00)	0.00(0.00)	0.2
Quantile	4.07(0.84)	4.13(0.61)	0.73	0.34(0.37)	0.25(0.42)	2.28	0.85(1.12)	0.50(0.65)	10.64
EQC	3.01(0.04)	3.00(0.00)	0.24	0.00(0.00)	0.00(0.00)	0.89	0.00(0.00)	0.00(0.00)	3.21
nBayes	37.64(3.99)	38.00(3.73)	0.01	33.21(2.70)	32.75(3.08)	0.26	30.54(1.75)	30.63(1.96)	1.11
LDA	17.66(2.00)	17.75(1.77)	0.01	8.83 (1.31)	8.88(1.21)	0.06	13.65(2.17)	13.50(2.10)	0.6
1-NN	29.37(2.17)	29.00(1.96)	0.01	22.80(1.99)	23.00 (2.19)	0.01	20.33 (3.53)	19.50 (2.47)	0.02
SVM	11.24(1.55)	11.25 (1.54)	0.01	0.28 (0.28)	0.25 (0.37)	0.03	0.00(0.00)	0.00(0.00)	0.14
AdaBoost	6.52(1.28)	6.50 (1.49)	4.39	0.95 (0.57)	1.00(0.56)	8.80	0.43(0.31)	0.50(0.23)	36.89
GBDT	4.98(1.05)	5.00(0.75)	3.89	0.20(0.23)	0.13(0.23)	7.21	0.00(0.00)	0.00(0.00)	20.46
XGBoost	32.08(2.23)	32.13(1.87)	0.01	31.31 (2.52)	31.25(2.10)	0.05	30.09(2.21)	30.00(1.91)	0.02
$n = 200$ (balanced)									
Unimodal(ROT)	10.53(0.77)	10.60(0.75)	0.86	1.55(0.28)	1.60(0.24)	3.28	0.00(0.00)	0.00(0.00)	14.04
Unimodal(NS)	5.78(0.54)	5.65(0.52)	0.88	0.15(0.10)	0.15(0.07)	2.56	0.00(0.00)	0.00(0.00)	10.78
Quantile-mode	7.90(0.63)	7.90(0.62)	8.00	0.48(0.18)	0.50(0.22)	8.96	0.00(0.00)	0.00(0.00)	11.05
Naive-mode	4.87(6.86)	4.70(6.72)	0.07	0.03(0.60)	0.00(0.00)	0.08	0.00(0.00)	0.00(0.00)	0.19
Centroid	13.61(0.74)	13.65(0.88)	0.01	4.77(0.61)	4.70(0.75)	0.01	24.26(24.65)	1.35(36.34)	0.01
Median	2.77(0.40)	2.75(0.37)	0.02	0.04(0.05)	0.00 (0.07)	0.08	0.00(0.00)	0.00(0.00)	0.24
Quantile	2.94(0.51)	2.80(0.39)	0.82	0.10 (0.15)	0.10 (0.07)	2.44	0.14(0.20)	0.10(0.15)	10.82
EQC	2.89(0.09)	2.90(0.00)	0.26	0.00(0.00)	0.00(0.00)	0.80	0.00(0.00)	0.00(0.00)	3.69
nBayes	39.39(4.16)	40.70(3.82)	0.13	35.54(3.25)	35.95 (3.21)	0.42	30.32(1.46)	30.10(1.42)	2.19
LDA	14.35(0.87)	14.50(0.84)	0.02	6.08(0.63)	6.10 (0.39)	0.08	3.77 (0.60)	3.70(0.62)	0.69
1-NN	25.50(1.12)	25.50(1.14)	0.01	23.37(1.69)	23.80 (1.98)	0.01	17.08(1.23)	17.05(1.18)	0.05
SVM	11.58(0.70)	11.60(0.63)	0.02	1.02(0.29)	1.00 (0.30)	0.04	0.00(0.00)	0.00(0.00)	0.24
AdaBoost	4.57(0.57)	4.60(0.63)	5.27	0.46(0.23)	0.40 (0.22)	11.14	0.13(0.10)	0.10(0.15)	49.21
GBDT	3.24(0.44)	3.20(0.45)	5.14	0.09(0.08)	0.10 (0.07)	8.45	0.00(0.00)	0.00(0.00)	38.14
XGBoost	30.84(1.57)	31.05(1.90)	0.01	31.84(1.60)	31.75 (1.90)	0.02	30.54(1.24)	30.55(0.11)	0.03

Table S2: Misclassification rates multiplied by 100 for Case 1 of Example 1 with imbalanced setting

	$p = 50$			$p = 200$			$p = 1000$		
	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time
$n = 50$ (imbalanced)									
Unimodal(ROT)	8.27(2.12)	8.00 (2.61)	0.67	0.45(0.49)	0.67(0.50)	3.19	0.00(0.00)	0.00 (0.00)	13.21
Unimodal(NS)	5.85(1.54)	6.00(1.49)	0.60	0.02(0.11)	0.00(0.00)	2.36	0.00(0.00)	0.00(0.00)	10.83
Quantile-mode	3.92(1.46)	3.33 (1.12)	4.50	0.00(0.00)	0.00(0.00)	8.10	0.00(0.00)	0.00(0.00)	9.70
Naive-mode	6.50(43.60)	6.17(3.73)	0.06	0.83(1.34)	0.83(1.49)	2.36	0.08(0.42)	0.00(0.00)	10.83
Centroid	19.67(2.41)	19.67(2.11)	0.01	8.28(2.23)	8.00(2.11)	0.01	6.03(4.69)	3.33 (6.47)	0.01
Median	3.18(1.23)	2.67(1.49)	0.02	0.00(0.00)	0.00(0.00)	0.07	0.00(0.00)	0.00(0.00)	0.22
Quantile	4.32(1.86)	4.00(1.49)	0.69	2.57(3.66)	1.33(1.99)	2.34	20.22(9.51)	22.33(11.19)	10.25
EQC	5.97(0.21)	6.00(0.00)	0.24	0.02(0.11)	0.00(0.00)	0.74	0.00(0.00)	0.00(0.00)	3.22
nBayes	36.10(3.73)	36.00(2.74)	0.06	32.12(3.22)	33.33(3.61)	0.19	30.00(2.88)	29.67 (2.49)	0.87
LDA	20.77(2.21)	20.67(2.11)	0.01	19.60(3.89)	19.33(3.48)	0.04	12.97(3.34)	13.33(3.11)	0.33
1-NN	28.78(2.50)	28.67(1.99)	0.01	22.07(3.92)	22.00(3.11)	0.01	23.02(5.26)	22.33 (6.47)	0.01
SVM	16.50(3.16)	16.00(2.99)	0.01	0.02(0.11)	0.00(0.00)	0.02	0.00(0.00)	0.00(0.00)	0.13
AdaBoost	9.78(2.04)	10.00(2.11)	4.31	4.88(2.09)	4.67(1.62)	0.74	4.87(2.15)	4.67(1.99)	3.22
GBDT	6.57(2.05)	6.70(1.62)	4.07	0.92(0.89)	0.67(1.00)	4.82	0.28(0.42)	0.00 (0.50)	18.60
XGBoost	28.68(3.60)	29.33(2.99)	0.01	27.82(2.95)	28.00(3.11)	0.74	29.65(3.08)	30.00(2.24)	3.22
$n = 100$ (imbalanced)									
Unimodal(ROT)	11.10 (0.95)	11.25 (1.03)	0.92	0.38 (0.21)	0.33 (0.12)	3.21	0.00 (0.00)	0.00 (0.00)	13.60
Unimodal(NS)	5.96 (1.03)	6.25 (0.90)	0.86	0.02 (0.05)	0.00 (0.00)	2.39	0.00 (0.00)	0.00 (0.00)	10.44
Quantile-mode	6.47 (0.71)	6.50 (0.78)	7.71	0.00 (0.00)	0.00 (0.00)	8.68	0.00 (0.00)	0.00 (0.00)	10.53
Naive-mode	5.09 (4.18)	5.25 (3.73)	0.06	0.17 (1.06)	0.17 (0.93)	0.08	0.00 (0.00)	0.00 (0.00)	0.17
Centroid	16.04 (1.18)	16.00 (0.87)	0.01	4.83 (0.69)	4.67 (0.62)	0.01	1.19 (0.48)	1.33 (0.56)	0.01
Median	2.51 (0.51)	2.42 (0.44)	0.02	0.00 (0.00)	0.00 (0.00)	0.07	0.00 (0.00)	0.00 (0.00)	0.23
Quantile	2.61 (0.59)	2.75 (0.65)	0.78	0.23 (0.26)	0.17 (0.25)	2.79	0.42 (0.65)	0.17 (0.37)	10.72
EQC	4.66(0.13)	4.67(0.00)	0.36	0.65(0.11)	0.67(0.00)	1.14	0.16(0.03)	0.17(0.00)	3.59
SVM	12.72 (1.14)	12.67 (1.03)	0.01	0.51 (0.33)	0.50 (0.37)	0.06	0.00 (0.00)	0.00 (0.00)	0.21
LDA	17.35 (1.18)	17.33 (1.40)	0.02	7.88 (1.03)	7.83 (0.81)	0.07	12.56 (2.27)	13.33 (2.61)	0.49
1-NN	24.34 (1.22)	24.58 (1.37)	0.01	19.06 (1.75)	18.92 (1.80)	0.01	16.64 (2.40)	16.67 (1.49)	0.03
nBayes	41.63 (6.79)	42.33 (8.55)	0.11	32.88 (3.13)	32.83 (3.36)	0.35	31.41 (1.94)	31.17 (1.49)	1.75
AdaBoost	5.90 (0.94)	5.83 (0.87)	4.98	1.38 (0.51)	1.33 (0.37)	9.98	0.99 (0.38)	1.08 (0.50)	41.05
GBDT	3.96 (0.78)	4.00 (0.65)	4.79	0.14 (0.15)	0.17 (0.16)	7.29	0.00 (0.00)	0.00 (0.00)	30.15
XGBoost	28.62 (1.53)	28.50 (1.52)	0.02	26.72 (1.71)	26.42 (1.27)	0.02	27.07 (1.99)	26.67 (2.08)	0.03
$n = 200$ (imbalanced)									
Unimodal(ROT)	12.25 (0.59)	12.27 (0.41)	0.97	1.11 (0.22)	1.13 (0.25)	3.37	0.07 (0.05)	0.07 (0.05)	14.96
Unimodal(NS)	5.87 (0.06)	5.90 (0.61)	0.90	0.08 (0.06)	0.07 (0.06)	2.62	0.00 (0.00)	0.00 (0.00)	11.79
Quantile-mode	8.73 (0.46)	8.63 (0.47)	9.88	0.30 (0.12)	0.30 (0.15)	10.00	0.00 (0.00)	0.00 (0.00)	12.34
Naive-mode	3.97 (8.02)	4.03 (9.14)	0.07	0.05 (0.71)	0.07 (0.75)	0.09	0.00 (0.00)	0.00 (0.00)	0.23
Centroid	15.33 (0.71)	15.17 (0.68)	0.01	4.98 (0.77)	5.17 (0.96)	0.01	1.25 (0.21)	1.20 (0.20)	0.02
Median	2.41 (0.31)	2.37 (0.36)	0.02	0.00 (0.00)	0.00 (0.00)	0.05	0.00 (0.00)	0.00 (0.00)	0.22
Quantile	2.66 (0.46)	2.60 (0.37)	0.95	0.06 (0.06)	0.07 (0.05)	2.60	0.08 (0.11)	0.07 (0.10)	12.02
EQC	2.98(0.15)	3.00(0.00)	0.31	0.07(0.01)	0.07(0.00)	0.93	0.00(0.00)	0.00(0.00)	4.36
SVM	13.17 (0.73)	13.10 (0.68)	0.03	1.44 (0.29)	1.47 (0.30)	0.06	0.00 (0.00)	0.00 (0.00)	0.32
LDA	15.50 (0.76)	15.47 (0.82)	0.02	5.20 (0.50)	5.07 (0.62)	0.10	2.46 (0.31)	2.53 (0.31)	1.08
1-NN	22.96 (0.86)	23.00 (0.76)	0.01	19.07 (1.04)	19.10 (0.91)	0.02	15.44 (0.97)	15.27 (0.61)	0.14
nBayes	34.07 (5.12)	32.33 (5.19)	0.18	29.87 (2.40)	29.63 (1.87)	0.67	29.96 (1.61)	29.80 (1.69)	3.46
AdaBoost	4.29 (0.52)	4.23 (0.56)	5.74	0.55 (0.16)	0.50 (0.20)	13.49	0.23 (0.12)	0.20 (0.10)	60.01
GBDT	3.25 (0.48)	3.13 (0.61)	5.94	0.07 (0.06)	0.07 (0.10)	20.87	0.00 (0.00)	0.00 (0.00)	53.38
XGBoost	26.40 (1.21)	26.30 (1.42)	0.02	26.06 (1.03)	25.97 (1.26)	0.93	25.55 (1.35)	25.27 (1.28)	4.36

Table S3: Misclassification rates multiplied by 100 for Case 2 of Example 1 with balanced setting

	$p = 50$			$p = 200$			$p = 1000$		
	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time
$n = 50$ (balanced)									
Unimodal(ROT)	0.00 (0.00)	0.00 (0.00)	0.96	0.00 (0.00)	0.00 (0.00)	2.61	0.00 (0.00)	0.00 (0.00)	10.84
Unimodal(NS)	0.03 (0.16)	0.00 (0.00)	0.87	0.00 (0.00)	0.00 (0.00)	2.24	0.00 (0.00)	0.00 (0.00)	8.66
Quantile-mode	0.00 (0.00)	0.00 (0.00)	1.99	0.00 (0.00)	0.00 (0.00)	2.11	0.00 (0.00)	0.00 (0.00)	9.53
Naive-mode	50.00 (0.00)	50.00 (0.00)	0.07	50.00 (0.00)	50.00 (0.00)	0.07	50.00 (0.00)	50.00 (0.00)	0.13
Centroid	3.15 (0.95)	3.00 (1.49)	0.01	0.00 (0.00)	0.00 (0.00)	0.01	0.00 (0.00)	0.00 (0.00)	0.01
Meidan	0.78 (0.83)	1.00 (0.75)	0.01	0.00 (0.00)	0.00 (0.00)	0.05	0.00 (0.00)	0.00 (0.00)	0.21
Quantile	0.35 (1.23)	0.00 (0.00)	0.64	0.03 (0.16)	0.00 (0.00)	2.20	0.00 (0.00)	0.00 (0.00)	9.88
EQC	0.00 (0.00)	0.00 (0.00)	0.24	0.00 (0.00)	0.00 (0.00)	0.73	0.00 (0.00)	0.00 (0.00)	3.01
nBayes	8.75 (2.96)	9.00 (2.43)	0.04	1.95 (1.50)	1.00 (1.49)	0.15	0.05 (0.22)	0.00 (0.00)	0.72
LDA	8.28 (2.89)	8.00 (3.17)	0.02	1.48 (1.32)	1.00 (0.93)	0.04	0.58 (0.81)	0.00 (0.75)	0.26
1-NN	17.65 (2.68)	18.00 (1.87)	0.01	6.33 (4.43)	5.50 (5.40)	0.01	0.00 (0.00)	0.00 (0.00)	0.01
SVM	2.78 (1.39)	2.50 (1.49)	0.01	0.00 (0.00)	0.00 (0.00)	0.02	0.00 (0.00)	0.00 (0.00)	0.15
AdaBoost	0.43 (0.64)	0.00 (0.75)	4.01	0.05 (0.22)	0.00 (0.00)	7.85	0.05 (0.22)	0.00 (0.00)	31.61
GBDT	0.10 (0.30)	0.00 (0.00)	4.15	0.00 (0.00)	0.00 (0.00)	4.98	0.00 (0.00)	0.00 (0.00)	18.53
XGBoost	15.48 (4.14)	15.00 (3.54)	0.01	16.33 (5.03)	16.00 (4.66)	0.02	16.85 (4.34)	16.50 (4.66)	0.02
$n = 100$ (balanced)									
Unimodal(ROT)	0.03 (0.08)	0.00 (0.00)	0.74	0.00 (0.00)	0.00 (0.00)	3.14	0.00 (0.00)	0.00 (0.00)	11.20
Unimodal(NS)	0.13 (0.13)	0.25 (0.19)	0.74	0.00 (0.00)	0.00 (0.00)	2.31	0.00 (0.00)	0.00 (0.00)	9.28
Quantile-mode	0.03 (0.08)	0.00 (0.00)	1.32	0.00 (0.00)	0.00 (0.00)	2.22	0.00 (0.00)	0.00 (0.00)	10.21
Naive-mode	50.00 (0.00)	50.00 (0.00)	0.06	50.00 (0.00)	50.00 (0.00)	0.08	50.00 (0.00)	50.00 (0.00)	0.16
Centroid	1.43 (0.60)	1.50 (0.42)	0.01	0.00 (0.00)	0.00 (0.00)	0.01	0.00 (0.00)	0.00 (0.00)	0.01
Median	0.76 (0.42)	0.75 (0.37)	0.03	0.00 (0.00)	0.00 (0.00)	0.06	0.00 (0.00)	0.00 (0.00)	0.23
Quantile	0.00 (0.00)	0.00 (0.00)	0.72	0.00 (0.00)	0.00 (0.00)	2.32	0.00 (0.00)	0.00 (0.00)	10.56
EQC	0.00 (0.00)	0.00 (0.00)	0.25	0.00 (0.00)	0.00 (0.00)	0.75	0.00 (0.00)	0.00 (0.00)	3.19
nBayes	4.51 (1.04)	4.50 (0.79)	0.07	0.10 (0.16)	0.00 (0.19)	0.29	0.00 (0.00)	0.00 (0.00)	1.33
LDA	2.28 (0.66)	2.25 (0.75)	0.01	0.29 (0.23)	0.25 (0.19)	0.06	0.48 (0.39)	0.50 (0.37)	0.41
1-NN	9.13 (1.59)	9.13 (1.40)	0.01	1.64 (0.83)	1.50 (0.56)	0.01	0.05 (0.12)	0.00 (0.00)	0.02
SVM	1.93 (0.98)	1.63 (1.03)	0.01	0.00 (0.00)	0.00 (0.00)	0.05	0.00 (0.00)	0.00 (0.00)	0.16
AdaBoost	0.03 (0.08)	0.00 (0.00)	4.30	0.00 (0.00)	0.00 (0.00)	8.60	0.00 (0.00)	0.00 (0.00)	34.81
GBDT	0.00 (0.00)	0.00 (0.00)	4.62	0.00 (0.00)	0.00 (0.00)	4.84	0.00 (0.00)	0.00 (0.00)	21.84
XGBoost	6.03 (1.34)	5.88 (0.98)	0.02	5.91 (1.62)	5.63 (0.75)	0.04	5.92 (1.33)	6.13 (1.54)	0.04
$n = 200$ (balanced)									
Unimodal(ROT)	0.00 (0.00)	0.00 (0.00)	0.84	0.00 (0.00)	0.00 (0.00)	2.91	0.00 (0.00)	0.00 (0.00)	13.20
Unimodal(NS)	0.00 (0.00)	0.00 (0.00)	0.89	0.00 (0.00)	0.00 (0.00)	2.23	0.00 (0.00)	0.00 (0.00)	10.39
Quantile-mode	0.00 (0.00)	0.00 (0.00)	2.30	0.00 (0.00)	0.00 (0.00)	2.44	0.00 (0.00)	0.00 (0.00)	11.24
Naïve-mode	50.00 (0.00)	50.00 (0.00)	0.07	50.00 (0.00)	50.00 (0.00)	0.09	50.00 (0.00)	50.00 (0.00)	0.19
Centroid	1.49 (0.29)	1.40 (0.30)	0.01	0.00 (0.00)	0.00 (0.00)	0.01	0.00 (0.00)	0.00 (0.00)	0.01
Median	1.22 (0.26)	1.20 (0.22)	0.02	0.00 (0.00)	0.00 (0.00)	0.08	0.00 (0.00)	0.00 (0.00)	0.26
Quantile	0.00 (0.00)	0.00 (0.00)	0.84	0.00 (0.00)	0.00 (0.00)	2.43	0.00 (0.00)	0.00 (0.00)	11.10
EQC	0.01 (0.02)	0.01 (0.00)	0.27	0.00 (0.00)	0.00 (0.00)	0.80	0.00 (0.00)	0.00 (0.00)	4.05
nBayes	3.00 (0.66)	2.80 (0.67)	0.13	0.01 (0.03)	0.00 (0.00)	0.44	0.00 (0.00)	0.00 (0.00)	2.25
LDA	1.96 (0.41)	1.85 (0.39)	0.02	0.01 (0.02)	0.00 (0.00)	0.09	0.00 (0.00)	0.00 (0.00)	0.63
1-NN	10.31 (1.10)	10.30 (1.42)	0.01	1.98 (0.57)	2.00 (0.56)	0.01	0.14 (0.12)	0.10 (0.15)	0.05
SVM	1.29 (0.28)	1.30 (0.30)	0.02	0.00 (0.00)	0.00 (0.00)	0.05	0.00 (0.00)	0.00 (0.00)	0.26
AdaBoost	0.00 (0.00)	0.00 (0.00)	4.66	0.00 (0.00)	0.00 (0.00)	9.89	0.00 (0.00)	0.00 (0.00)	49.79
GBDT	0.00 (0.00)	0.00 (0.00)	5.47	0.00 (0.00)	0.00 (0.00)	8.36	0.00 (0.00)	0.00 (0.00)	38.19
XGBoost	4.95 (0.69)	4.95 (0.76)	0.01	5.02 (0.79)	5.05 (0.49)	0.02	5.02 (0.73)	5.00 (0.94)	0.03

Table S4: Misclassification rates multiplied by 100 for Case 2 of Example 1 with imbalanced setting

	p = 50			p = 200			p = 1000		
	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time
$n = 50$ (imbalanced)									
Unimodal(ROT)	0.00 (0.00)	0.00 (0.00)	0.79	0.00 (0.00)	0.00 (0.00)	2.67	0.00 (0.00)	0.00 (0.00)	11.63
Unimodal(NS)	0.00 (0.00)	0.00 (0.00)	0.69	0.00 (0.00)	0.00 (0.00)	2.17	0.00 (0.00)	0.00 (0.00)	10.65
Quantile-mode	0.00 (0.00)	0.00 (0.00)	2.11	0.00 (0.00)	0.00 (0.00)	3.11	0.00 (0.00)	0.00 (0.00)	10.19
Naïve-mode	33.33 (0.00)	33.33 (0.00)	0.06	33.33 (0.00)	33.33 (0.00)	0.07	33.33 (0.00)	33.33 (0.00)	0.14
Centroid	1.35 (0.61)	1.33 (0.62)	0.01	0.00 (0.00)	0.00 (0.00)	0.01	0.00 (0.00)	0.00 (0.00)	0.01
Median	1.57 (0.83)	1.33 (1.00)	0.01	0.00 (0.00)	0.00 (0.00)	0.05	0.00 (0.00)	0.00 (0.00)	0.21
Quantile	0.12 (0.73)	0.00 (0.00)	0.71	0.00 (0.00)	0.00 (0.00)	2.23	0.00 (0.00)	0.00 (0.00)	10.14
EQC	0.00 (0.00)	0.00 (0.00)	0.25	0.00 (0.00)	0.00 (0.00)	0.73	0.00 (0.00)	0.00 (0.00)	3.17
SVM	33.33 (0.00)	33.33 (0.00)	0.01	33.33 (0.00)	33.33 (0.00)	0.02	10.65 (3.18)	10.67 (3.61)	0.13
LDA	3.88 (1.43)	3.33 (1.24)	0.02	3.42 (1.93)	3.33 (2.11)	0.04	0.23 (0.39)	0.00 (0.50)	0.35
1-NN	12.15 (1.76)	12.33 (1.62)	0.01	1.67 (0.81)	2.00 (0.62)	0.01	0.78 (0.93)	0.33 (1.00)	0.01
nBayes	7.70 (2.11)	7.67 (1.62)	0.05	1.67 (1.30)	1.33 (1.00)	0.22	0.18 (0.30)	0.00 (0.50)	0.95
AdaBoost	0.27 (0.39)	0.00 (0.50)	4.27	0.17 (0.33)	0.00 (0.00)	8.27	0.10 (0.24)	0.00 (0.00)	33.55
GBDT	0.00 (0.00)	0.00 (0.00)	4.30	0.00 (0.00)	0.00 (0.00)	4.54	0.00 (0.00)	0.00 (0.00)	18.18
XGBoost	9.25 (2.61)	8.67 (2.49)	0.01	11.60 (3.35)	11.00 (2.49)	0.02	10.70 (2.67)	10.33 (2.11)	0.02
$n = 100$ (imbalanced)									
Unimodal(ROT)	0.00 (0.00)	0.00 (0.00)	0.86	0.00 (0.00)	0.00 (0.00)	2.75	0.00 (0.00)	0.00 (0.00)	11.20
Unimodal(NS)	0.01 (0.04)	0.00 (0.00)	0.74	0.00 (0.00)	0.00 (0.00)	2.15	0.00 (0.00)	0.00 (0.00)	9.31
Quantile-mode	0.00 (0.00)	0.00 (0.00)	2.51	0.00 (0.00)	0.00 (0.00)	5.63	0.00 (0.00)	0.00 (0.00)	10.70
Naive-mode	33.33 (0.00)	33.33 (0.00)	0.07	33.33 (0.00)	33.33 (0.00)	0.08	33.33 (0.00)	33.33 (0.00)	0.17
Centroid	1.55 (0.42)	1.50 (0.28)	0.01	0.00 (0.00)	0.00 (0.00)	0.01	0.00 (0.00)	0.00 (0.00)	0.01
Median	1.51 (0.40)	1.50 (0.28)	0.02	0.00 (0.00)	0.00 (0.00)	0.06	0.00 (0.00)	0.00 (0.00)	0.22
Quantile	0.00 (0.00)	0.00 (0.00)	0.79	0.00 (0.00)	0.00 (0.00)	2.63	0.00 (0.00)	0.00 (0.00)	10.55
EQC	0.00 (0.00)	0.00 (0.00)	0.27	0.00 (0.00)	0.00 (0.00)	0.83	0.00 (0.00)	0.00 (0.00)	3.72
SVM	33.33 (0.00)	33.33 (0.00)	0.01	0.19 (0.17)	0.17 (0.25)	0.03	0.00 (0.00)	0.00 (0.00)	0.21
LDA	1.89 (0.45)	2.00 (0.40)	0.01	0.10 (0.10)	0.17 (0.12)	0.06	0.28 (0.23)	0.17 (0.25)	0.48
1-NN	9.99 (1.11)	10.08 (0.93)	0.01	2.31 (0.71)	2.17 (0.87)	0.01	0.21 (0.18)	0.17 (0.25)	0.03
nBayes	3.13 (0.85)	3.00 (0.90)	0.10	0.13 (0.13)	0.17 (0.12)	0.36	0.00 (0.00)	0.00 (0.00)	1.73
AdaBoost	0.05 (0.08)	0.00 (0.12)	4.73	0.00 (0.00)	0.00 (0.00)	9.23	0.00 (0.00)	0.00 (0.00)	40.33
GBDT	0.00 (0.00)	0.00 (0.00)	5.26	0.00 (0.00)	0.00 (0.00)	6.58	0.00 (0.00)	0.00 (0.00)	28.56
XGBoost	3.52 (0.74)	3.50 (0.65)	0.01	3.98 (1.06)	3.83 (0.78)	0.02	4.26 (1.38)	4.17 (0.78)	0.03
$n = 200$ (imbalanced)									
Unimodal(ROT)	0.00 (0.01)	0.00 (0.00)	1.13	0.00 (0.00)	0.00 (0.00)	5.03	0.00 (0.00)	0.00 (0.00)	13.99
Unimodal(NS)	0.00 (0.00)	0.00 (0.00)	1.34	0.00 (0.00)	0.00 (0.00)	3.72	0.00 (0.00)	0.00 (0.00)	12.38
Quantile-mode	0.00 (0.01)	0.00 (0.00)	2.50	0.00 (0.00)	0.00 (0.00)	6.21	0.00 (0.00)	0.00 (0.00)	12.75
Naive-mode	33.33 (0.00)	33.33 (0.00)	0.07	33.33 (0.00)	33.33 (0.00)	0.09	33.33 (0.00)	33.33 (0.00)	0.23
Centroid	1.28 (0.22)	1.27 (0.21)	0.01	0.00 (0.00)	0.00 (0.00)	0.01	0.00 (0.00)	0.00 (0.00)	0.02
Median	0.77 (0.20)	0.73 (0.21)	0.02	0.00 (0.00)	0.00 (0.00)	0.05	0.00 (0.00)	0.00 (0.00)	0.25
Quantile	0.00 (0.00)	0.00 (0.00)	0.98	0.00 (0.00)	0.00 (0.00)	3.21	0.00 (0.00)	0.00 (0.00)	11.95
EQC	0.00 (0.02)	0.00 (0.00)	0.32	0.00 (0.00)	0.00 (0.00)	0.96	0.00 (0.00)	0.00 (0.00)	4.46
SVM	26.09 (1.30)	26.17 (1.33)	0.03	0.00 (0.00)	0.00 (0.00)	0.07	0.00 (0.00)	0.00 (0.00)	0.38
LDA	1.41 (0.28)	1.33 (0.31)	0.02	0.00 (0.00)	0.00 (0.00)	0.10	0.00 (0.00)	0.00 (0.00)	1.37
1-NN	8.68 (0.67)	8.70 (0.81)	0.01	1.63 (0.32)	1.67 (0.26)	0.02	0.06 (0.06)	0.07 (0.05)	0.12
nBayes	2.07 (0.48)	2.03 (0.44)	0.18	0.05 (0.05)	0.07 (0.05)	0.69	0.00 (0.00)	0.00 (0.00)	3.33
AdaBoost	0.00 (0.00)	0.00 (0.00)	5.26	0.00 (0.00)	0.00 (0.00)	11.96	0.00 (0.00)	0.00 (0.00)	52.03
GBDT	0.00 (0.00)	0.00 (0.00)	6.34	0.00 (0.00)	0.00 (0.00)	20.67	0.00 (0.00)	0.00 (0.00)	52.55
XGBoost	3.32 (0.37)	3.37 (0.32)	0.02	3.19 (0.47)	3.13 (0.42)	0.02	3.17 (0.50)	3.10 (0.35)	0.04

Table S5: Misclassification rates multiplied by 100 for Case 3 of Example 1 with balanced setting

	$p = 50$			$p = 200$			$p = 1000$		
	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time
$n = 50$ (balanced)									
Unimodal(ROT)	13.60 (3.47)	13.50 (3.92)	2.07	4.83 (2.06)	5.00 (1.87)	4.44	0.10 (0.30)	0.00 (0.00)	11.05
Unimodal(NS)	15.63 (2.62)	15.00 (2.99)	1.50	9.95 (2.59)	9.50 (2.99)	3.73	1.23 (0.89)	1.00 (0.75)	9.61
Quantile-mode	11.60 (2.70)	12.00 (2.43)	2.03	2.88 (1.26)	3.00 (0.93)	8.04	0.00 (0.00)	0.00 (0.00)	38.83
Naive-mode	30.68 (5.10)	31.00 (5.97)	0.06	32.23 (8.71)	32.50 (11.38)	0.08	10.83 (7.71)	9.50 (8.02)	0.13
Centroid	49.00 (3.09)	50.00 (0.75)	0.01	48.98 (2.15)	50.00 (0.75)	0.01	49.93 (0.47)	50.00 (0.00)	0.01
Median	18.68 (3.37)	18.50 (2.43)	0.02	5.90 (1.97)	6.00 (1.68)	0.06	0.23 (0.48)	0.00 (0.00)	0.25
Quantile	8.73 (3.45)	8.00 (2.99)	0.65	2.70 (2.48)	2.00 (2.24)	2.15	1.08 (1.65)	0.00 (1.49)	9.56
EQC	18.08 (0.47)	18.00 (0.00)	0.23	2.03 (0.16)	2.00 (0.00)	0.72	0.00 (0.00)	0.00 (0.00)	2.98
nBayes	12.83 (2.71)	13.00 (2.43)	0.04	7.33 (3.10)	7.50 (2.99)	0.13	2.65 (1.56)	3.00 (2.24)	0.67
LDA	27.83 (4.99)	27.00 (4.85)	0.02	26.98 (6.15)	27.00 (5.41)	0.04	21.08 (5.19)	21.00 (4.66)	0.30
1-NN	10.73 (2.55)	10.50 (2.24)	0.01	15.63 (4.24)	15.00 (4.66)	0.02	12.15 (3.40)	12.50 (3.73)	0.12
SVM	2.88 (1.52)	3.00 (1.49)	0.01	0.00 (0.00)	0.00 (0.00)	0.02	0.00 (0.00)	0.00 (0.00)	0.12
AdaBoost	3.88 (1.96)	4.00 (2.24)	4.50	1.53 (1.15)	1.00 (0.75)	8.05	0.38 (0.54)	0.00 (0.75)	31.65
GBDT	4.20 (1.96)	4.00 (2.24)	4.42	1.55 (0.88)	2.00 (0.75)	5.06	0.38 (0.90)	0.00 (0.19)	20.03
XGBoost	22.48 (5.35)	23.00 (6.34)	0.01	22.05 (4.55)	21.50 (5.41)	0.02	22.00 (4.45)	21.00 (4.48)	0.02
$n = 100$ (balanced)									
Unimodal(ROT)	12.83 (1.81)	12.75 (2.10)	1.78	2.93 (0.74)	3.00 (0.75)	3.57	0.01 (0.04)	0.00 (0.00)	12.28
Unimodal(NS)	15.33 (1.79)	15.50 (1.96)	1.61	8.33 (1.47)	8.25 (1.31)	2.89	0.17 (0.17)	0.25 (0.19)	10.08
Quantile-mode	8.57 (1.04)	8.50 (1.12)	2.27	1.00 (0.33)	1.00 (0.37)	8.36	0.00 (0.00)	0.00 (0.00)	41.10
Naive-mode	49.56 (0.46)	49.75 (0.23)	0.07	45.52 (3.40)	46.25 (2.85)	0.07	5.64 (2.45)	5.50 (1.96)	0.15
Centroid	49.45 (1.97)	50.00 (0.19)	0.01	49.74 (0.33)	49.75 (0.19)	0.01	49.99 (0.25)	50.00 (0.00)	0.01
Median	16.93 (1.67)	16.50 (1.63)	0.02	2.97 (0.69)	2.88 (0.56)	0.05	0.19 (0.19)	0.25 (0.19)	0.23
Quantile	7.53 (1.71)	7.00 (1.40)	0.73	0.11 (0.17)	0.00 (0.19)	2.19	0.00 (0.00)	0.00 (0.00)	10.16
EQC	15.74 (0.08)	15.75 (0.00)	0.25	1.98 (0.12)	2.00 (0.00)	0.69	0.00 (0.00)	0.00 (0.00)	3.18
nBayes	18.56 (3.01)	18.88 (2.71)	0.06	4.67 (1.14)	4.50 (0.93)	0.23	1.56 (0.55)	1.63 (0.61)	1.20
LDA	27.63 (2.47)	27.25 (2.19)	0.01	17.63 (2.59)	17.75 (2.85)	0.04	23.34 (3.01)	23.50 (3.08)	0.38
1-NN	14.44 (1.54)	14.50 (1.49)	0.01	7.46 (1.37)	7.75 (1.49)	0.01	9.09 (1.39)	9.13 (1.54)	0.02
SVM	5.73 (0.93)	5.75 (0.98)	0.01	0.00 (0.00)	0.00 (0.00)	0.04	0.01 (0.04)	0.00 (0.00)	0.13
AdaBoost	2.89 (0.83)	2.75 (0.93)	4.28	0.03 (0.08)	0.00 (0.00)	8.84	0.00 (0.00)	0.00 (0.00)	35.28
GBDT	2.69 (0.76)	2.50 (0.98)	4.66	0.02 (0.07)	0.00 (0.00)	4.17	0.00 (0.00)	0.00 (0.00)	21.36
XGBoost	19.87 (4.06)	20.75 (5.64)	0.01	15.14 (2.91)	14.88 (1.96)	0.02	14.24 (2.01)	14.38 (1.91)	0.02
$n = 200$ (balanced)									
Unimodal(ROT)	15.25 (1.89)	14.60 (2.85)	1.76	2.09 (0.47)	2.05 (0.54)	3.22	0.01 (0.02)	0.00 (0.00)	13.91
Unimodal(NS)	20.28 (2.05)	20.60 (2.72)	1.62	2.86 (0.67)	2.75 (0.71)	2.98	0.36 (0.21)	0.30 (0.22)	10.18
Quantile-mode	10.57 (0.73)	10.55 (0.97)	5.34	0.61 (0.19)	0.60 (0.17)	20.10	0.00 (0.00)	0.00 (0.00)	45.34
Naïve-mode	32.30 (3.15)	32.10 (2.93)	0.11	13.34 (2.33)	12.85 (1.87)	0.64	7.98 (2.89)	7.30 (2.78)	2.12
Centroid	48.81 (1.22)	49.30 (0.76)	0.01	49.76 (0.29)	49.80 (0.15)	0.01	49.99 (0.04)	50.00 (0.00)	0.01
Median	20.18 (0.84)	20.15 (0.99)	0.02	2.03 (0.34)	2.00 (0.37)	0.05	0.25 (0.13)	0.20 (0.07)	0.23
Quantile	7.85 (0.59)	7.85 (0.63)	0.78	0.08 (0.08)	0.10 (0.07)	2.52	0.00 (0.00)	0.00 (0.00)	10.62
EQC	13.99 (0.09)	14.00 (0.00)	0.30	0.99 (0.05)	1.00 (0.00)	0.84	0.10 (0.02)	0.10 (0.00)	3.56
nBayes	21.21 (1.63)	21.10 (1.53)	0.11	3.50 (0.50)	3.45 (0.45)	0.64	1.01 (0.35)	0.95 (0.32)	2.12
LDA	34.35 (2.05)	34.25 (1.92)	0.02	14.09 (1.31)	14.05 (1.19)	0.10	9.25 (1.31)	9.20 (1.12)	0.63
1-NN	15.12 (0.79)	15.20 (0.75)	0.01	3.60 (0.70)	3.55 (0.67)	0.01	8.26 (0.79)	8.35 (0.69)	0.05
SVM	7.45 (0.56)	7.50 (0.39)	0.01	0.00 (0.00)	0.00 (0.00)	0.03	0.09 (0.46)	0.00 (0.00)	0.17
AdaBoost	2.10 (0.44)	2.20 (0.52)	5.41	0.01 (0.03)	0.00 (0.00)	11.24	0.00 (0.00)	0.00 (0.00)	46.08
GBDT	1.84 (0.40)	1.90 (0.45)	5.11	0.00 (0.00)	0.00 (0.00)	9.06	0.00 (0.00)	0.00 (0.00)	38.37
XGBoost	16.59 (3.83)	14.75 (4.85)	0.30	14.07 (2.18)	13.90 (1.14)	0.84	13.45 (1.24)	13.50 (1.32)	3.56

Table S6: Misclassification rates multiplied by 100 for Case 3 of Example 1 with imbalanced setting

	$p = 50$			$p = 200$			$p = 1000$		
	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time
$n = 50$ (imbalanced)									
Unimodal(ROT)	10.45 (1.81)	10.67 (2.49)	1.88	6.62 (1.78)	7.00 (1.49)	4.35	0.35 (0.43)	0.00 (0.50)	11.80
Unimodal(NS)	14.55 (2.44)	14.00 (2.49)	1.66	11.08 (2.20)	10.67 (1.99)	3.01	1.15 (0.75)	1.33 (0.62)	9.76
Quantile-mode	10.97 (1.99)	10.67 (1.49)	8.63	2.55 (1.12)	2.67 (1.00)	18.61	0.00 (0.00)	0.00 (0.00)	40.25
Naive-mode	65.55 (2.02)	66.00 (1.00)	0.06	15.30 (5.87)	13.67 (6.59)	0.08	21.05 (10.45)	20.33 (8.21)	0.16
Centroid	62.97 (7.16)	65.00 (2.24)	0.01	40.32 (13.42)	33.33 (0.00)	0.01	48.63 (16.54)	33.33 (24.38)	0.01
Median	18.00 (3.09)	17.67 (3.48)	0.01	7.50 (1.52)	7.33 (1.49)	0.04	0.80 (0.63)	0.67 (0.62)	0.22
Quantile	8.90 (2.89)	8.00 (2.24)	0.70	0.82 (0.94)	0.67 (1.00)	2.55	0.78 (0.95)	0.67 (0.62)	9.81
EQC	18.58 (0.53)	18.67 (0.00)	0.25	0.67 (0.00)	0.67 (0.00)	0.69	0.70 (0.21)	0.67 (0.00)	3.28
SVM	4.43 (1.50)	4.33 (1.99)	0.01	0.00 (0.00)	0.00 (0.00)	0.03	0.00 (0.00)	0.00 (0.00)	0.14
LDA	18.27 (3.35)	18.00 (2.49)	0.02	26.68 (5.30)	27.33 (4.60)	0.06	18.15 (2.98)	18.00 (2.11)	0.35
1-NN	21.33 (2.75)	21.33 (2.61)	0.01	8.73 (2.97)	9.00 (3.61)	0.01	9.42 (2.33)	9.00 (1.99)	0.01
nBayes	18.33 (4.39)	18.00 (4.23)	0.06	8.10 (2.31)	8.33 (2.99)	0.26	2.22 (1.42)	2.00 (1.12)	0.94
AdaBoost	4.37 (1.35)	4.67 (1.49)	4.66	1.20 (0.76)	1.33 (0.50)	8.31	0.38 (0.50)	0.00 (0.50)	34.94
GBDT	5.97 (2.29)	5.33 (1.87)	4.01	0.52 (0.55)	0.67 (0.50)	4.74	0.00 (0.18)	0.00 (0.00)	18.74
XGBoost	20.55 (3.63)	20.00 (2.49)	0.01	12.73 (4.18)	12.00 (4.10)	0.02	16.62 (3.38)	16.67 (3.98)	0.02
$n = 100$ (imbalanced)									
Unimodal(ROT)	12.93 (1.42)	13.00 (1.15)	1.75	0.93 (0.39)	0.83 (0.37)	4.01	0.09 (0.12)	0.00 (0.12)	10.89
Unimodal(NS)	19.19 (1.31)	19.17 (1.31)	1.98	3.59 (0.76)	3.67 (0.78)	3.24	0.74 (0.32)	0.67 (0.37)	9.71
Quantile-mode	12.00 (0.99)	12.00 (1.12)	32.04	0.77 (0.30)	0.67 (0.25)	36.15	0.01 (0.04)	0.00 (0.00)	43.45
Naive-mode	29.71 (1.71)	29.33 (1.40)	0.08	22.83 (6.38)	22.33 (7.46)	0.09	5.75 (2.71)	4.92 (2.77)	0.19
Centroid	63.13 (5.25)	64.33 (1.06)	0.01	39.85 (13.09)	33.30 (0.12)	0.01	62.93 (10.00)	66.33 (0.50)	0.01
Median	18.26 (1.69)	18.17 (1.87)	0.02	4.06 (0.67)	4.17 (0.65)	0.08	0.21 (0.17)	0.17 (0.16)	0.24
Quantile	5.29 (0.84)	5.42 (0.87)	0.78	0.09 (0.10)	0.00 (0.12)	2.67	0.00 (0.00)	0.00 (0.00)	10.38
EQC	10.00 (0.00)	10.00 (0.00)	0.27	1.83 (0.03)	1.83 (0.00)	0.79	0.33 (0.05)	0.33 (0.00)	3.72
SVM	6.20 (0.99)	6.33 (1.03)	0.01	0.00 (0.00)	0.00 (0.00)	0.03	0.00 (0.00)	0.00 (0.00)	0.15
LDA	26.48 (1.37)	26.42 (1.49)	0.02	12.08 (1.34)	12.00 (1.52)	0.07	21.43 (2.25)	21.08 (2.30)	0.53
1-NN	16.46 (1.22)	16.67 (1.06)	0.01	9.12 (1.01)	9.00 (0.93)	0.01	4.84 (0.83)	4.75 (0.78)	0.03
nBayes	29.68 (2.26)	29.58 (2.52)	0.10	3.52 (0.54)	3.50 (0.50)	0.36	0.68 (0.32)	0.67 (0.25)	1.67
AdaBoost	2.20 (0.63)	2.25 (0.50)	5.05	0.00 (0.03)	0.00 (0.00)	9.59	0.00 (0.03)	0.00 (0.00)	40.86
GBDT	2.56 (0.85)	2.33 (0.68)	5.27	0.24 (0.24)	0.17 (0.25)	6.99	0.00 (0.00)	0.00 (0.00)	29.96
XGBoost	18.17 (1.78)	17.75 (0.78)	0.02	10.38 (1.23)	10.25 (1.15)	0.02	12.22 (1.39)	12.25 (1.55)	0.03
$n = 200$ (imbalanced)									
Unimodal(ROT)	14.23 (1.57)	14.13 (1.26)	1.91	4.48 (0.62)	4.50 (0.52)	3.99	0.06 (0.05)	0.07 (0.05)	19.59
Unimodal(NS)	18.37 (1.20)	18.30 (1.02)	1.85	7.49 (0.71)	7.53 (0.63)	3.31	1.07 (0.26)	1.07 (0.35)	14.64
Quantile-mode	10.94 (0.57)	10.93 (0.51)	2.91	0.78 (0.17)	0.73 (0.15)	10.69	0.00 (0.00)	0.00 (0.00)	49.84
Naive-mode	48.63 (4.68)	48.83 (4.88)	0.07	14.24 (2.54)	14.17 (2.15)	0.11	5.63 (1.72)	5.93 (1.85)	0.29
Centroid	43.91 (15.79)	33.33 (24.78)	0.01	64.06 (8.91)	66.60 (0.15)	0.01	65.58 (5.27)	66.47 (0.11)	0.02
Median	14.90 (0.06)	14.87 (0.48)	0.02	3.68 (0.32)	3.67 (0.34)	0.05	0.19 (0.09)	0.20 (0.10)	0.28
Quantile	4.69 (0.56)	4.57 (0.56)	0.93	0.09 (0.10)	0.07 (0.10)	2.55	0.00 (0.00)	0.00 (0.00)	11.83
EQC	11.65 (0.08)	11.67 (0.00)	0.30	0.99 (0.05)	1.00 (0.00)	0.89	0.00 (0.00)	0.00 (0.00)	4.31
SVM	10.82 (16.98)	4.93 (0.67)	0.02	7.54 (12.40)	0.20 (13.48)	0.05	0.09 (0.45)	0.00 (0.00)	0.26
LDA	22.58 (1.05)	22.67 (0.98)	0.02	11.35 (1.07)	11.37 (0.71)	0.10	6.51 (0.81)	6.40 (0.95)	1.00
1-NN	14.02 (0.87)	14.17 (0.75)	0.01	14.83 (0.95)	14.63 (0.90)	0.02	4.57 (0.50)	4.53 (0.51)	0.12
nBayes	10.90 (1.38)	10.83 (0.97)	0.17	5.95 (1.11)	5.90 (1.07)	0.63	0.85 (0.23)	0.87 (0.22)	3.21
AdaBoost	1.17 (0.27)	1.13 (0.31)	5.99	0.02 (0.04)	0.00 (0.00)	13.32	0.00 (0.00)	0.00 (0.00)	56.74
GBDT	1.35 (0.36)	1.27 (0.32)	6.31	0.05 (0.06)	0.07 (0.05)	21.31	0.00 (0.00)	0.00 (0.00)	54.70
XGBoost	14.07 (1.20)	14.17 (1.39)	0.02	11.42 (0.84)	11.43 (0.86)	0.02	10.60 (0.97)	10.60 (0.82)	0.04

Table S7: Misclassification rates multiplied by 100 for Case 4 of Example 1 with balanced setting

	$p = 50$			$p = 200$			$p = 1000$		
	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time
$n = 50$ (balanced)									
Unimodal(ROT)	0.45 (0.68)	0.00 (0.75)	1.01	0.00 (0.00)	0.00 (0.00)	2.04	0.00 (0.00)	0.00 (0.00)	9.22
Unimodal(NS)	1.73 (1.13)	2.00 (0.75)	0.82	0.00 (0.00)	0.00 (0.00)	2.01	0.00 (0.00)	0.00 (0.00)	8.46
Quantile-mode	0.00 (0.00)	0.00 (0.00)	2.28	0.00 (0.00)	0.00 (0.00)	3.23	0.00 (0.00)	0.00 (0.00)	10.84
Naive-mode	48.30 (2.02)	49.00 (2.43)	0.07	49.18 (1.28)	50.00 (0.75)	0.10	50.00 (0.00)	50.00 (0.00)	0.18
Centroid	8.48 (2.66)	9.00 (2.99)	0.01	2.30 (1.24)	2.00 (1.49)	0.01	0.20 (0.41)	0.00 (0.00)	0.01
Median	0.85 (0.66)	1.00 (0.75)	0.01	0.00 (0.00)	0.00 (0.00)	0.06	0.00 (0.00)	0.00 (0.00)	0.18
Quantile	0.55 (0.64)	0.00 (0.75)	0.59	0.05 (0.22)	0.00 (0.00)	2.12	0.00 (0.00)	0.00 (0.00)	9.48
EQC	0.00 (0.00)	0.00 (0.00)	0.20	0.00 (0.00)	0.00 (0.00)	0.65	0.00 (0.00)	0.00 (0.00)	2.91
nBayes	5.33 (2.22)	5.00 (1.49)	0.04	3.03 (1.51)	3.00 (1.49)	0.13	1.88 (1.42)	2.00 (1.49)	0.66
LDA	25.93 (7.97)	26.00 (7.84)	0.01	0.43 (0.75)	0.00 (0.75)	0.13	0.13 (0.40)	0.00 (0.00)	0.66
1-NN	24.20 (4.26)	25.00 (2.99)	0.01	19.95 (4.43)	21.00 (5.22)	0.01	10.83 (3.97)	10.00 (4.48)	0.01
SVM	3.25 (1.32)	3.50 (1.49)	0.01	0.00 (0.00)	0.00 (0.00)	0.02	0.00 (0.00)	0.00 (0.00)	0.13
AdaBoost	0.25 (0.49)	0.00 (0.00)	3.85	0.00 (0.00)	0.00 (0.00)	7.63	0.00 (0.00)	0.00 (0.00)	30.66
GBDT	0.05 (0.22)	0.00 (0.00)	3.42	0.00 (0.00)	0.00 (0.00)	4.02	0.00 (0.00)	0.00 (0.00)	18.69
XGBoost	13.35 (4.49)	13.50 (5.41)	0.01	14.28 (3.84)	14.50 (4.48)	0.02	15.28 (4.95)	14.50 (5.22)	0.02
$n = 100$ (balanced)									
Unimodal(ROT)	0.19 (0.20)	0.25 (0.19)	0.97	0.00 (0.00)	0.00 (0.00)	1.97	0.00 (0.00)	0.00 (0.00)	9.78
Unimodal(NS)	1.86 (0.66)	1.88 (0.75)	0.85	0.08 (0.12)	0.00 (0.19)	1.85	0.00 (0.00)	0.00 (0.00)	8.57
Quantile-mode	0.04 (0.10)	0.00 (0.00)	2.81	0.00 (0.00)	0.00 (0.00)	2.91	0.00 (0.00)	0.00 (0.00)	12.40
Naive-mode	49.98 (0.07)	50.00 (0.00)	0.06	50.00 (0.00)	50.00 (0.00)	0.25	50.00 (0.00)	50.00 (0.00)	1.12
Centroid	2.94 (0.88)	2.75 (0.93)	0.01	1.03 (0.45)	0.10 (0.37)	0.01	0.09 (0.14)	0.00 (0.19)	0.01
Median	0.22 (0.19)	0.25 (0.19)	0.02	0.00 (0.00)	0.00 (0.00)	0.04	0.00 (0.00)	0.00 (0.00)	0.19
Quantile	0.16 (0.26)	0.00 (0.19)	0.67	0.00 (0.00)	0.00 (0.00)	2.13	0.00 (0.00)	0.00 (0.00)	9.66
EQC	0.00 (0.02)	0.00 (0.00)	0.22	0.00 (0.00)	0.00 (0.00)	0.69	0.00 (0.00)	0.00 (0.00)	3.12
nBayes	1.23 (0.64)	1.25 (0.56)	0.06	0.70 (0.41)	0.50 (0.37)	0.25	0.46 (0.24)	0.50 (0.23)	1.12
LDA	0.08 (0.15)	0.00 (0.00)	0.01	0.03 (0.08)	0.00 (0.00)	0.05	0.06 (0.14)	0.00 (0.00)	0.37
1-NN	17.31 (2.03)	17.13 (1.91)	0.01	15.69 (1.69)	15.50 (1.87)	0.01	8.06 (1.28)	8.00 (1.31)	0.02
SVM	0.46 (0.29)	0.50 (0.37)	0.01	0.00 (0.00)	0.00 (0.00)	0.03	0.00 (0.00)	0.00 (0.00)	0.14
AdaBoost	0.00 (0.00)	0.00 (0.00)	4.18	0.00 (0.00)	0.00 (0.00)	7.99	0.00 (0.00)	0.00 (0.00)	34.07
GBDT	0.00 (0.00)	0.00 (0.00)	3.80	0.00 (0.00)	0.00 (0.00)	4.75	0.00 (0.00)	0.00 (0.00)	21.17
XGBoost	3.49 (1.85)	3.00 (0.93)	0.01	3.46 (1.60)	3.00 (1.40)	0.02	4.37 (1.62)	4.63 (1.59)	0.02
$n = 200$ (balanced)									
Unimodal(ROT)	0.48 (0.16)	0.50 (0.15)	0.87	0.01 (0.02)	0.00 (0.00)	2.47	0.00 (0.00)	0.00 (0.00)	10.43
Unimodal(NS)	2.44 (0.46)	2.40 (0.39)	0.93	0.30 (0.14)	0.30 (0.15)	2.39	0.00 (0.00)	0.00 (0.00)	10.28
Quantile-mode	0.17 (0.12)	0.15 (0.15)	3.47	0.00 (0.00)	0.00 (0.00)	8.08	0.00 (0.00)	0.00 (0.00)	15.49
Naive-mode	50.00 (0.00)	50.00 (0.00)	0.08	50.00 (0.00)	50.00 (0.00)	0.13	50.00 (0.00)	50.00 (0.00)	0.39
Centroid	3.73 (0.59)	3.80 (0.67)	0.01	0.60 (0.17)	0.60 (0.15)	0.01	0.04 (0.05)	0.00 (0.07)	0.01
Median	0.22 (0.11)	0.20 (0.09)	0.02	0.00 (0.00)	0.00 (0.00)	0.05	0.00 (0.00)	0.00 (0.00)	0.20
Quantile	0.04 (0.07)	0.00 (0.00)	0.76	0.00 (0.02)	0.00 (0.00)	2.43	0.00 (0.00)	0.00 (0.00)	10.33
EQC	0.00 (0.00)	0.00 (0.00)	0.25	0.00 (0.00)	0.00 (0.00)	0.77	0.00 (0.00)	0.00 (0.00)	3.47
nBayes	0.50 (0.25)	0.50 (0.30)	0.12	0.38 (0.18)	0.40 (0.22)	0.44	0.15 (0.12)	0.20 (0.09)	2.13
LDA	0.00 (0.02)	0.00 (0.00)	0.02	0.01 (0.02)	0.00 (0.00)	0.09	0.00 (0.00)	0.00 (0.00)	0.60
1-NN	19.00 (1.33)	19.05 (1.53)	0.01	13.18 (0.76)	13.20 (0.84)	0.01	8.89 (1.16)	8.90 (1.23)	0.05
SVM	0.37 (0.13)	0.40 (0.08)	0.02	0.01 (0.03)	0.00 (0.00)	0.04	0.00 (0.00)	0.00 (0.00)	0.20
AdaBoost	0.00 (0.00)	0.00 (0.00)	4.95	0.00 (0.00)	0.00 (0.00)	9.51	0.00 (0.00)	0.00 (0.00)	42.04
GBDT	0.00 (0.00)	0.00 (0.00)	4.57	0.00 (0.00)	0.00 (0.00)	8.61	0.00 (0.00)	0.00 (0.00)	35.18
XGBoost	3.01 (0.46)	2.85 (0.52)	0.02	2.87 (0.51)	2.85 (0.54)	0.02	2.83 (5.00)	2.75 (0.56)	0.03

Table S8: Misclassification rates multiplied by 100 for Case 4 of Example 1 with imbalanced setting

	$p = 50$			$p = 200$			$p = 1000$		
	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time
$n = 50$ (imbalanced)									
Unimodal(ROT)	0.43 (0.44)	0.67 (0.50)	0.80	0.00 (0.00)	0.00 (0.00)	2.02	0.00 (0.00)	0.00 (0.00)	9.06
Unimodal(NS)	1.40 (0.88)	1.33 (1.00)	0.93	0.03 (0.15)	0.00 (0.00)	1.99	0.00 (0.00)	0.00 (0.00)	8.45
Quantile-mode	0.08 (0.22)	0.00 (0.00)	1.98	0.00 (0.00)	0.00 (0.00)	4.79	0.00 (0.00)	0.00 (0.00)	11.63
Naive-mode	32.92 (0.58)	33.33 (0.50)	0.07	33.33 (0.00)	33.33 (0.00)	0.10	33.33 (0.00)	33.33 (0.00)	0.22
Centroid	5.32 (1.73)	4.67 (1.99)	0.01	1.47 (0.98)	1.33 (1.00)	0.01	0.00 (0.00)	0.00 (0.00)	0.01
Median	0.02 (0.11)	0.00 (0.00)	0.01	0.02 (0.11)	0.00 (0.00)	0.04	0.00 (0.00)	0.00 (0.00)	0.20
Quantile	0.18 (0.45)	0.00 (0.00)	0.62	0.02 (0.11)	0.00 (0.00)	2.21	0.00 (0.00)	0.00 (0.00)	9.61
EQC	0.67 (0.00)	0.67 (0.00)	0.21	0.00 (0.00)	0.00 (0.00)	0.67	0.00 (0.00)	0.00 (0.00)	3.18
SVM	7.20 (2.46)	7.33 (2.74)	0.01	0.08 (0.22)	0.00 (0.00)	0.02	0.00 (0.00)	0.00 (0.00)	0.15
LDA	1.65 (1.56)	1.33 (1.00)	0.01	1.18 (0.87)	1.33 (0.50)	0.04	0.02 (0.11)	0.00 (0.00)	0.30
1-NN	20.22 (2.60)	20.00 (2.49)	0.01	15.47 (2.68)	15.00 (2.61)	0.01	13.42 (2.91)	13.33 (2.99)	0.01
nBayes	2.72 (1.34)	3.00 (1.12)	0.04	2.88 (1.21)	2.67 (1.00)	0.19	1.03 (0.72)	0.67 (0.50)	1.14
AdaBoost	0.07 (0.20)	0.00 (0.00)	3.98	0.03 (0.15)	0.00 (0.00)	7.69	0.00 (0.00)	0.00 (0.00)	32.19
GBDT	0.00 (0.00)	0.00 (0.00)	3.65	0.00 (0.00)	0.00 (0.00)	4.25	0.00 (0.00)	0.00 (0.00)	18.25
XGBoost	6.85 (1.87)	7.33 (1.62)	0.01	8.60 (3.18)	8.33 (3.48)	0.02	8.93 (2.19)	8.67 (1.99)	0.02
$n = 100$ (imbalanced)									
Unimodal(ROT)	0.51 (0.26)	0.50 (0.25)	1.16	0.00 (0.00)	0.00 (0.00)	2.16	0.00 (0.00)	0.00 (0.00)	10.05
Unimodal(NS)	2.47 (0.48)	2.50 (0.50)	1.30	0.12 (0.13)	0.17 (0.12)	2.12	0.00 (0.00)	0.00 (0.00)	10.13
Quantile-mode	0.03 (0.06)	0.00 (0.00)	2.71	0.00 (0.00)	0.00 (0.00)	4.27	0.00 (0.00)	0.00 (0.00)	13.88
Naive-mode	33.33 (0.04)	33.33 (0.00)	0.09	33.33 (0.00)	33.33 (0.00)	0.33	33.33 (0.00)	33.33 (0.00)	1.65
Centroid	3.75 (0.65)	3.83 (0.41)	0.01	0.76 (0.30)	0.67 (0.37)	0.01	0.05 (0.09)	0.00 (0.00)	0.01
Median	0.10 (0.10)	0.17 (0.12)	0.02	0.00 (0.00)	0.00 (0.00)	0.05	0.00 (0.00)	0.00 (0.00)	0.19
Quantile	0.10 (0.15)	0.00 (0.00)	0.71	0.00 (0.00)	0.00 (0.00)	2.30	0.00 (0.00)	0.00 (0.00)	10.03
EQC	0.00 (0.00)	0.00 (0.00)	0.23	0.00 (0.00)	0.00 (0.00)	0.77	0.00 (0.00)	0.00 (0.00)	3.45
SVM	0.65 (0.26)	0.67 (0.25)	0.01	0.00 (0.03)	0.00 (0.00)	0.03	0.00 (0.00)	0.00 (0.00)	0.16
LDA	0.02 (0.05)	0.00 (0.00)	0.01	0.00 (0.00)	0.00 (0.00)	0.06	0.05 (0.10)	0.00 (0.12)	0.50
1-NN	16.85 (1.16)	17.00 (1.15)	0.01	14.40 (1.50)	14.17 (1.27)	0.01	10.25 (2.55)	9.25 (2.67)	0.03
nBayes	0.76 (0.26)	0.67 (0.37)	0.09	0.54 (0.24)	0.50 (0.25)	0.33	0.23 (0.21)	0.17 (0.25)	1.65
AdaBoost	0.00 (0.00)	0.00 (0.00)	4.48	0.00 (0.00)	0.00 (0.00)	8.90	0.00 (0.00)	0.00 (0.00)	38.04
GBDT	0.00 (0.00)	0.00 (0.00)	4.62	0.00 (0.00)	0.00 (0.00)	6.42	0.00 (0.00)	0.00 (0.00)	28.59
XGBoost	2.63 (0.95)	2.33 (0.75)	0.01	3.18 (1.31)	3.25 (1.68)	0.02	3.27 (0.96)	3.25 (0.62)	0.03
$n = 200$ (imbalanced)									
Unimodal(ROT)	0.57 (0.15)	0.60 (0.15)	1.23	0.02 (0.03)	0.00 (0.05)	2.66	0.00 (0.00)	0.00 (0.00)	11.36
Unimodal(NS)	2.32 (0.37)	2.27 (0.36)	1.03	0.17 (0.11)	0.13 (0.15)	2.11	0.00 (0.00)	0.00 (0.00)	10.40
Quantile-mode	0.32 (0.08)	0.33 (0.05)	2.57	0.00 (0.00)	0.00 (0.00)	4.08	0.00 (0.00)	0.00 (0.00)	18.85
Naive-mode	33.33 (0.00)	33.33 (0.00)	0.15	33.33 (0.00)	33.33 (0.00)	0.61	33.33 (0.00)	33.33 (0.00)	2.97
Centroid	3.38 (0.31)	3.33 (0.31)	0.01	0.69 (0.16)	0.67 (0.15)	0.02	0.04 (0.04)	0.03 (0.05)	0.02
Median	0.15 (0.06)	0.13 (0.05)	0.02	0.00 (0.00)	0.00 (0.00)	0.07	0.00 (0.00)	0.00 (0.00)	0.22
Quantile	0.05 (0.11)	0.00 (0.05)	0.84	0.00 (0.01)	0.00 (0.00)	2.73	0.00 (0.00)	0.00 (0.00)	11.32
EQC	0.00 (0.00)	0.00 (0.00)	0.28	0.00 (0.00)	0.00 (0.00)	0.95	0.00 (0.00)	0.00 (0.00)	4.21
SVM	0.35 (0.13)	0.33 (0.11)	0.02	0.00 (0.00)	0.00 (0.00)	0.08	0.00 (0.00)	0.00 (0.00)	0.32
LDA	0.03 (0.03)	0.00 (0.05)	0.02	0.00 (0.00)	0.00 (0.00)	0.12	0.00 (0.00)	0.00 (0.00)	0.96
1-NN	15.04 (0.89)	15.03 (0.72)	0.01	14.12 (0.85)	14.10 (0.76)	0.02	10.35 (1.17)	10.43 (1.08)	0.12
nBayes	0.66 (0.21)	0.60 (0.21)	0.15	0.40 (0.16)	0.40 (0.15)	0.61	0.05 (0.07)	0.00 (0.05)	2.97
AdaBoost	0.00 (0.00)	0.00 (0.00)	4.96	0.00 (0.00)	0.00 (0.00)	11.97	0.00 (0.00)	0.00 (0.00)	51.73
GBDT	0.00 (0.00)	0.00 (0.00)	5.76	0.00 (0.00)	0.00 (0.00)	20.01	0.00 (0.00)	0.00 (0.00)	49.95
XGBoost	1.35 (0.38)	1.33 (0.41)	0.01	2.01 (0.34)	2.00 (0.35)	0.02	1.70 (0.41)	1.60 (0.25)	0.04

Table S9: Misclassification rates multiplied by 100 for Case 5 of Example 1 with balanced setting and $\rho = 0.2$

$n = 50$ (balanced)									
	$p = 50$			$p = 200$			$p = 1000$		
	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time
Unimodal(ROT)	0.00 (0.00)	0.00 (0.00)	1.03	0.00 (0.00)	0.00 (0.00)	2.08	0.00 (0.00)	0.00 (0.00)	8.79
Unimodal(NS)	0.65 (0.80)	0.00 (0.75)	1.04	0.00 (0.00)	0.00 (0.00)	1.93	0.00 (0.00)	0.00 (0.00)	8.09
Quantile-mode	0.00 (0.00)	0.00 (0.00)	2.35	0.00 (0.00)	0.00 (0.00)	5.74	0.00 (0.00)	0.00 (0.00)	10.90
Naive-mode	48.68 (2.02)	50.00 (1.68)	0.07	49.44 (1.11)	50.00 (0.75)	0.08	50.00 (0.00)	50.00 (0.00)	0.18
Centroid	6.88 (2.48)	7.00 (1.87)	0.01	3.19 (1.31)	3.00 (1.49)	0.01	0.18 (0.38)	0.00 (0.00)	0.01
Median	0.20 (0.41)	0.00 (0.00)	0.01	0.00 (0.00)	0.00 (0.00)	0.04	0.00 (0.00)	0.00 (0.00)	0.19
Quantile	0.68 (0.83)	0.00 (0.75)	0.61	0.03 (0.17)	0.00 (0.00)	2.18	0.00 (0.00)	0.00 (0.00)	9.40
nBayes	4.83 (1.80)	4.50 (2.24)	0.03	4.33 (2.03)	5.00 (2.24)	0.13	2.15 (1.58)	2.00 (1.49)	0.66
EQC	0.00 (0.00)	0.00 (0.00)	0.21	0.00 (0.00)	0.00 (0.00)	0.63	0.00 (0.00)	0.00 (0.00)	2.88
LDA	25.98 (8.47)	24.00 (10.63)	0.01	0.58 (0.75)	0.00 (0.75)	0.03	0.13 (0.40)	0.00 (0.00)	0.26
1-NN	23.35 (3.70)	22.00 (4.48)	0.01	18.33 (3.50)	19.00 (3.73)	0.01	15.78 (3.60)	16.00 (3.17)	0.01
SVM	1.75 (1.35)	2.00 (1.49)	0.01	0.06 (0.24)	0.00 (0.00)	0.02	0.00 (0.00)	0.00 (0.00)	0.11
AdaBoost	0.03 (0.16)	0.00 (0.00)	3.94	0.00 (0.00)	0.00 (0.00)	7.66	0.00 (0.00)	0.00 (0.00)	31.06
GBDT	0.00 (0.00)	0.00 (0.00)	3.58	0.00 (0.00)	0.00 (0.00)	4.34	0.00 (0.00)	0.00 (0.00)	19.23
XGBoost	16.65 (4.87)	18.00 (5.97)	0.01	14.88 (3.91)	15.00 (3.92)	0.02	16.23 (4.97)	16.00 (5.97)	0.02
$n = 100$ (balanced)									
Unimodal(ROT)	0.43 (0.27)	0.50 (0.37)	1.09	0.00 (0.00)	0.00 (0.00)	2.10	0.00 (0.00)	0.00 (0.00)	9.19
Unimodal(NS)	2.38 (0.57)	2.25 (0.56)	0.96	0.16 (0.17)	0.25 (0.19)	1.88	0.00 (0.00)	0.00 (0.00)	8.70
Quantile-mode	0.00 (0.00)	0.00 (0.00)	2.86	0.00 (0.00)	0.00 (0.00)	6.58	0.00 (0.00)	0.00 (0.00)	12.41
Naïve-mode	50.00 (0.00)	50.00 (0.00)	0.07	50.00 (0.00)	50.00 (0.00)	0.10	50.00 (0.00)	50.00 (0.00)	0.26
Centroid	5.33 (1.02)	5.38 (0.75)	0.01	1.96 (0.65)	2.00 (0.56)	0.01	0.28 (0.24)	0.25 (0.37)	0.01
Median	0.31 (0.20)	0.25 (0.19)	0.01	0.00 (0.00)	0.00 (0.00)	0.04	0.00 (0.00)	0.00 (0.00)	0.21
Quantile	0.14 (0.21)	0.00 (0.19)	0.66	0.01 (0.04)	0.00 (0.00)	2.25	0.00 (0.00)	0.00 (0.00)	9.74
EQC	0.00 (0.00)	0.00 (0.00)	0.22	0.00 (0.00)	0.00 (0.00)	0.76	0.00 (0.00)	0.00 (0.00)	3.08
nBayes	1.13 (0.58)	1.13 (0.56)	0.06	0.38 (0.34)	0.38 (0.37)	0.24	0.49 (0.30)	0.50 (0.37)	1.11
LDA	0.09 (0.17)	0.00 (0.19)	0.01	0.01 (0.04)	0.00 (0.00)	0.05	0.09 (0.18)	0.00 (0.05)	0.37
1-NN	23.73 (2.01)	24.00 (1.40)	0.01	16.34 (1.50)	16.13 (1.73)	0.01	12.01 (2.44)	12.00 (3.03)	0.02
SVM	1.13 (0.45)	1.00 (0.56)	0.01	0.08 (0.13)	0.00 (0.19)	0.03	0.00 (0.00)	0.00 (0.00)	0.14
AdaBoost	0.00 (0.00)	0.00 (0.00)	4.17	0.00 (0.00)	0.00 (0.00)	8.48	0.00 (0.00)	0.00 (0.00)	34.16
GBDT	0.00 (0.00)	0.00 (0.00)	3.80	0.00 (0.00)	0.00 (0.00)	4.16	0.00 (0.00)	0.00 (0.00)	20.40
XGBoost	2.88 (1.10)	2.75 (0.93)	0.01	4.32 (2.93)	3.50 (1.91)	0.02	4.22 (1.67)	3.75 (1.49)	0.02
$n = 200$ (balanced)									
Unimodal(ROT)	0.58 (0.19)	0.60 (0.09)	0.98	0.00 (0.00)	0.00 (0.00)	2.20	0.00 (0.00)	0.00 (0.00)	10.38
Unimodal(NS)	2.50 (0.44)	2.40 (0.37)	0.92	0.17 (0.11)	0.15 (0.07)	2.07	0.00 (0.00)	0.00 (0.00)	9.81
Quantile-mode	0.34 (0.12)	0.30 (0.07)	3.20	0.00 (0.00)	0.00 (0.00)	6.17	0.00 (0.00)	0.00 (0.00)	15.84
Naïve-mode	49.96 (0.05)	50.00 (0.07)	0.07	50.00 (0.00)	50.00 (0.00)	0.12	50.00 (0.00)	50.00 (0.00)	0.37
Centroid	5.59 (0.66)	5.55 (0.54)	0.01	1.59 (0.37)	1.50 (0.48)	0.01	0.01 (0.03)	0.00 (0.00)	0.01
Median	0.33 (0.14)	0.30 (0.15)	0.02	0.03 (0.04)	0.00 (0.07)	0.06	0.00 (0.00)	0.00 (0.00)	0.21
Quantile	0.03 (0.07)	0.00 (0.02)	0.77	0.00 (0.00)	0.00 (0.00)	2.39	0.00 (0.00)	0.00 (0.00)	10.72
EQC	0.00 (0.00)	0.00 (0.00)	0.25	0.00 (0.00)	0.00 (0.00)	0.80	0.00 (0.00)	0.00 (0.00)	3.46
nBayes	0.69 (0.35)	0.60 (0.26)	0.37	0.41 (0.19)	0.50 (0.17)	0.41	0.18 (0.12)	0.20 (0.09)	2.21
LDA	0.01 (0.03)	0.00 (0.00)	0.02	0.00 (0.00)	0.00 (0.00)	0.07	0.00 (0.00)	0.00 (0.00)	0.75
1-NN	18.31 (1.22)	18.30 (1.12)	0.01	16.46 (1.46)	16.50 (0.97)	0.01	9.01 (0.97)	8.95 (1.03)	0.05
SVM	0.52 (0.17)	0.50 (0.15)	0.02	0.01 (0.02)	0.00 (0.00)	0.03	0.00 (0.00)	0.00 (0.00)	0.30
AdaBoost	0.00 (0.00)	0.00 (0.00)	4.43	0.00 (0.00)	0.00 (0.00)	9.92	0.00 (0.00)	0.00 (0.00)	41.90
GBDT	0.00 (0.00)	0.00 (0.00)	5.06	0.00 (0.00)	0.00 (0.00)	7.84	0.00 (0.00)	0.00 (0.00)	39.15
XGBoost	2.88 (0.65)	2.90 (0.67)	0.01	2.77 (0.49)	2.80 (0.54)	0.02	2.72 (0.46)	2.70 (0.47)	0.03

Table S10: Misclassification rates multiplied by 100 for Case 5 of Example 1 with imbalanced setting and $\rho = 0.2$

	$p = 50$			$p = 200$			$p = 1000$		
	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time
$n = 50$ (imbalanced)									
Unimodal(ROT)	0.35 (0.37)	0.33 (0.50)	1.06	0.00 (0.00)	0.00 (0.00)	2.06	0.00 (0.00)	0.00 (0.00)	9.46
Unimodal(NS)	1.48 (0.87)	1.33 (1.00)	0.84	0.05 (0.18)	0.00 (0.00)	2.12	0.00 (0.00)	0.00 (0.00)	9.59
Quantile-mode	0.05 (0.18)	0.00 (0.00)	3.51	0.00 (0.00)	0.00 (0.00)	6.33	0.00 (0.00)	0.00 (0.00)	11.57
Naive-mode	32.88 (1.06)	33.33 (0.12)	0.07	33.28 (0.23)	33.33 (0.00)	0.09	33.33 (0.00)	33.33 (0.00)	0.22
Centroid	5.20 (1.61)	5.33 (1.00)	0.01	1.07 (0.62)	1.33 (0.50)	0.01	0.32 (0.43)	0.00 (0.50)	0.01
Median	0.17 (0.33)	0.00 (0.00)	0.03	0.00 (0.00)	0.00 (0.00)	0.04	0.00 (0.00)	0.00 (0.00)	0.20
Quantile	0.08 (0.22)	0.00 (0.00)	0.64	0.10 (0.24)	0.00 (0.00)	2.20	0.00 (0.00)	0.00 (0.00)	9.70
EQC	0.00 (0.00)	0.00 (0.00)	0.21	0.00 (0.00)	0.00 (0.00)	0.67	0.00 (0.00)	0.00 (0.00)	3.24
SVM	9.22 (2.85)	9.33 (2.99)	0.01	0.10 (0.28)	0.00 (0.00)	0.02	0.00 (0.00)	0.00 (0.00)	0.12
LDA	2.25 (1.03)	2.00 (1.12)	0.01	1.27 (1.18)	1.33 (0.50)	0.04	0.02 (0.11)	0.00 (0.00)	0.30
1-NN	20.45 (2.65)	20.67 (1.99)	0.01	18.05 (2.66)	18.33 (2.99)	0.01	9.53 (2.18)	10.00 (2.61)	0.01
nBayes	3.50 (1.69)	3.33 (1.99)	0.05	1.12 (0.65)	1.33 (0.50)	0.19	1.72 (1.06)	1.33 (1.12)	0.85
AdaBoost	0.15 (0.32)	0.00 (0.00)	3.99	0.02 (0.11)	0.00 (0.00)	7.71	0.02 (0.11)	0.00 (0.00)	32.32
GBDT	0.00 (0.00)	0.00 (0.00)	4.17	0.00 (0.00)	0.00 (0.00)	4.47	0.00 (0.00)	0.00 (0.00)	18.04
XGBoost	7.07 (1.97)	7.33 (1.62)	0.01	7.83 (2.64)	8.00 (1.99)	0.02	8.50 (2.34)	8.00 (1.62)	0.02
$n = 100$ (imbalanced)									
Unimodal(ROT)	0.47 (0.24)	0.50 (0.16)	1.06	0.00 (0.00)	0.00 (0.00)	2.32	0.00 (0.00)	0.00 (0.00)	9.91
Unimodal(NS)	2.00 (0.59)	2.00 (0.56)	1.19	0.06 (0.10)	0.00 (0.12)	2.03	0.00 (0.00)	0.00 (0.00)	9.56
Quantile-mode	0.10 (0.08)	0.17 (0.12)	4.47	0.00 (0.00)	0.00 (0.00)	7.36	0.00 (0.00)	0.00 (0.00)	13.91
Naive-mode	33.19 (0.27)	33.33 (0.12)	0.08	33.33 (0.00)	33.33 (0.00)	0.12	33.33 (0.00)	33.33 (0.00)	0.30
Centroid	4.50 (0.78)	4.50 (0.78)	0.01	1.33 (0.45)	1.17 (0.50)	0.01	0.07 (0.09)	0.00 (0.12)	0.01
Median	0.10 (0.08)	0.17 (0.12)	0.02	0.07 (0.08)	0.00 (0.12)	0.05	0.00 (0.00)	0.00 (0.00)	0.19
Quantile	0.12 (0.17)	0.00 (0.12)	0.69	0.00 (0.03)	0.00 (0.00)	2.36	0.00 (0.00)	0.00 (0.00)	10.05
EQC	0.00 (0.00)	0.00 (0.00)	0.23	0.00 (0.00)	0.00 (0.00)	0.75	0.00 (0.00)	0.00 (0.00)	3.45
SVM	0.91 (0.31)	0.92 (0.37)	0.01	0.06 (0.09)	0.00 (0.12)	0.03	0.00 (0.00)	0.00 (0.00)	0.19
LDA	0.03 (0.06)	0.00 (0.00)	0.01	0.00 (0.00)	0.00 (0.00)	0.06	0.06 (0.10)	0.00 (0.12)	0.47
1-NN	17.37 (1.40)	17.33 (1.05)	0.01	13.73 (1.27)	13.42 (1.49)	0.01	11.65 (1.49)	11.58 (1.37)	0.03
nBayes	1.28 (0.43)	1.33 (0.40)	0.09	0.58 (0.24)	0.67 (0.19)	0.32	0.20 (0.17)	0.17 (0.25)	1.66
AdaBoost	0.00 (0.00)	0.00 (0.00)	4.28	0.00 (0.00)	0.00 (0.00)	9.33	0.00 (0.00)	0.00 (0.00)	38.01
GBDT	0.00 (0.00)	0.00 (0.00)	4.74	0.00 (0.00)	0.00 (0.00)	6.16	0.00 (0.00)	0.00 (0.00)	28.19
XGBoost	2.64 (0.87)	2.50 (0.78)	0.01	3.26 (1.27)	3.83 (1.77)	0.02	3.80 (0.96)	4.08 (0.59)	0.03
$n = 200$ (imbalanced)									
Unimodal(ROT)	0.51 (0.18)	0.50 (0.20)	0.96	0.00 (0.01)	0.00 (0.00)	2.75	0.00 (0.00)	0.00 (0.00)	11.77
Unimodal(NS)	3.19 (0.51)	3.27 (0.51)	0.96	0.14 (0.09)	0.13 (0.10)	1.86	0.00 (0.00)	0.00 (0.00)	10.97
Quantile-mode	0.10 (0.06)	0.07 (0.05)	6.00	0.00 (0.00)	0.00 (0.00)	7.79	0.00 (0.00)	0.00 (0.00)	18.85
Naive-mode	33.33 (0.00)	33.33 (0.00)	0.11	33.33 (0.00)	33.33 (0.00)	0.16	33.33 (0.00)	33.33 (0.00)	0.51
Centroid	4.87 (0.50)	4.83 (0.46)	0.01	1.08 (0.23)	1.10 (0.20)	0.01	0.08 (0.05)	0.07 (0.05)	0.02
Median	0.13 (0.07)	0.13 (0.06)	0.02	0.00 (0.00)	0.00 (0.00)	0.06	0.00 (0.00)	0.00 (0.00)	0.22
Quantile	0.02 (0.05)	0.00 (0.00)	0.87	0.01 (0.02)	0.00 (0.00)	2.45	0.00 (0.00)	0.00 (0.00)	10.93
EQC	0.00 (0.00)	0.00 (0.00)	0.28	0.00 (0.00)	0.00 (0.00)	0.93	0.00 (0.00)	0.00 (0.00)	4.24
SVM	0.35 (0.12)	0.33 (0.10)	0.02	0.00 (0.01)	0.00 (0.00)	0.05	0.00 (0.00)	0.00 (0.00)	0.32
LDA	0.00 (0.00)	0.00 (0.00)	0.02	0.00 (0.00)	0.00 (0.00)	0.10	0.00 (0.00)	0.00 (0.00)	0.98
1-NN	17.16 (0.76)	17.27 (0.66)	0.01	15.72 (0.90)	15.60 (0.91)	0.02	9.60 (1.18)	9.53 (1.14)	0.11
nBayes	0.65 (0.19)	0.67 (0.15)	0.16	0.32 (0.14)	0.33 (0.11)	0.57	0.17 (0.08)	0.20 (0.11)	3.19
AdaBoost	0.00 (0.00)	0.00 (0.00)	5.10	0.00 (0.00)	0.00 (0.00)	11.62	0.00 (0.00)	0.00 (0.00)	51.98
GBDT	0.00 (0.00)	0.00 (0.00)	5.71	0.00 (0.00)	0.00 (0.00)	20.14	0.00 (0.00)	0.00 (0.00)	51.64
XGBoost	2.00 (0.37)	1.90 (0.36)	0.01	1.89 (0.38)	1.90 (0.26)	0.02	2.00 (0.39)	1.97 (0.35)	0.04

Table S11: Misclassification rates multiplied by 100 for Case 5 of Example 1 with balanced setting and $\rho = 0.5$

	$p = 50$			$p = 200$			$p = 1000$		
	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time
$n = 50$ (balanced)									
Unimodal(ROT)	0.20 (0.41)	0.00 (0.00)	1.04	0.00 (0.00)	0.00 (0.00)	1.99	0.00 (0.00)	0.00 (0.00)	8.98
Unimodal(NS)	1.53 (0.88)	1.00 (0.75)	0.93	0.00 (0.00)	0.00 (0.00)	1.97	0.00 (0.00)	0.00 (0.00)	8.42
Quantile-mode	0.03 (0.16)	0.00 (0.00)	3.32	0.00 (0.00)	0.00 (0.00)	5.69	0.00 (0.00)	0.00 (0.00)	10.66
Naive-mode	45.93 (3.97)	47.00 (3.73)	0.06	49.75 (0.63)	50.00 (0.00)	0.08	50.00 (0.00)	50.00 (0.00)	0.17
Centroid	11.93 (2.96)	12.00 (2.43)	0.01	7.18 (1.87)	7.00 (1.49)	0.01	1.48 (0.96)	1.00 (0.75)	0.01
Median	1.65 (1.14)	2.00 (0.93)	0.03	0.00 (0.00)	0.00 (0.00)	0.05	0.00 (0.00)	0.00 (0.00)	0.19
Quantile	1.13 (1.24)	1.00 (1.49)	0.61	0.00 (0.00)	0.00 (0.00)	2.22	0.00 (0.00)	0.00 (0.00)	9.29
EQC	0.00 (0.00)	0.00 (0.00)	0.21	0.00 (0.00)	0.00 (0.00)	0.65	0.00 (0.00)	0.00 (0.00)	2.85
nBayes	5.85 (2.62)	6.00 (2.24)	0.03	2.33 (1.02)	2.00 (0.75)	0.14	1.50 (1.22)	1.00 (0.75)	0.66
LDA	26.50 (7.47)	27.00 (7.65)	0.01	0.38 (0.74)	0.00 (0.19)	0.03	0.18 (0.38)	0.00 (0.00)	0.27
1-NN	29.43 (5.18)	29.50 (5.22)	0.01	21.30 (3.41)	21.00 (2.43)	0.01	20.23 (6.51)	18.50 (6.16)	0.01
SVM	5.15 (1.81)	5.00 (1.49)	0.01	0.58 (0.68)	0.00 (0.75)	0.02	0.00 (0.00)	0.00 (0.00)	0.11
AdaBoost	0.28 (0.55)	0.00 (0.00)	3.93	0.05 (0.22)	0.00 (0.00)	7.79	0.00 (0.00)	0.00 (0.00)	30.44
GBDT	0.00 (0.00)	0.00 (0.00)	3.60	0.00 (0.00)	0.00 (0.00)	4.16	0.00 (0.00)	0.00 (0.00)	19.97
XGBoost	14.18 (3.87)	14.00 (2.99)	0.01	15.75 (3.87)	15.00 (3.73)	0.02	16.90 (4.27)	16.00 (4.10)	0.02
$n = 100$ (balanced)									
Unimodal(ROT)	0.80 (0.37)	0.75 (0.37)	1.05	0.00 (0.00)	0.00 (0.00)	2.15	0.00 (0.00)	0.00 (0.00)	9.36
Unimodal(NS)	2.60 (0.66)	2.50 (0.79)	0.89	0.15 (0.16)	0.25 (0.19)	2.19	0.00 (0.00)	0.00 (0.00)	8.56
Quantile-mode	0.56 (0.29)	0.50 (0.23)	3.82	0.00 (0.00)	0.00 (0.00)	6.61	0.00 (0.00)	0.00 (0.00)	12.29
Naïve-mode	49.84 (0.23)	50.00 (0.19)	0.07	50.00 (0.00)	50.00 (0.00)	0.11	50.00 (0.00)	50.00 (0.00)	0.25
Centroid	9.61 (1.40)	9.75 (1.49)	0.01	4.40 (0.87)	4.50 (0.75)	0.01	0.61 (0.33)	0.50 (0.23)	0.01
Median	1.75 (0.52)	1.75 (0.37)	0.01	0.02 (0.07)	0.00 (0.00)	0.04	0.00 (0.00)	0.00 (0.00)	0.20
Quantile	0.20 (0.21)	0.25 (0.19)	0.67	0.05 (0.20)	0.00 (0.00)	2.14	0.00 (0.00)	0.00 (0.00)	9.98
EQC	0.00 (0.00)	0.00 (0.00)	0.25	0.00 (0.00)	0.00 (0.00)	0.72	0.00 (0.00)	0.00 (0.00)	3.09
nBayes	1.88 (0.66)	1.75 (0.75)	0.06	0.64 (0.41)	0.50 (0.51)	0.24	0.59 (0.31)	0.50 (0.19)	1.11
LDA	0.27 (0.21)	0.25 (0.37)	0.01	0.18 (0.23)	0.13 (0.19)	0.05	0.22 (0.27)	0.25 (0.23)	0.37
1-NN	21.21 (1.68)	20.88 (1.91)	0.01	20.61 (1.33)	20.38 (1.17)	0.01	16.73 (2.54)	17.50 (2.71)	0.01
SVM	2.50 (0.59)	2.50 (0.61)	0.01	0.23 (0.18)	0.25 (0.14)	0.03	0.00 (0.00)	0.00 (0.00)	0.14
AdaBoost	0.00 (0.00)	0.00 (0.00)	4.33	0.00 (0.00)	0.00 (0.00)	8.04	0.00 (0.00)	0.00 (0.00)	33.50
GBDT	0.00 (0.00)	0.00 (0.00)	3.75	0.00 (0.00)	0.00 (0.00)	4.82	0.00 (0.00)	0.00 (0.00)	21.97
XGBoost	3.70 (1.86)	3.25 (1.40)	0.01	3.62 (1.73)	3.00 (1.87)	0.01	4.79 (2.72)	1.41 (2.24)	0.02
$n = 200$ (balanced)									
Unimodal(ROT)	0.70 (0.23)	0.70 (0.30)	1.09	0.00 (0.02)	0.00 (0.00)	2.33	0.00 (0.00)	0.00 (0.00)	10.43
Unimodal(NS)	2.98 (0.51)	2.90 (0.52)	1.10	0.36 (0.15)	0.40 (0.15)	2.04	0.00 (0.00)	0.00 (0.00)	9.72
Quantile-mode	0.32 (0.14)	0.30 (0.15)	4.77	0.00 (0.00)	0.00 (0.00)	8.16	0.00 (0.00)	0.00 (0.00)	15.57
Naïve-mode	49.94 (0.09)	50.00 (0.07)	0.08	50.00 (0.00)	50.00 (0.00)	0.12	50.00 (0.00)	50.00 (0.00)	0.37
Centroid	9.90 (0.87)	9.95 (0.63)	0.01	4.11 (0.63)	4.05 (0.49)	0.01	0.30 (0.15)	0.30 (0.15)	0.01
Median	1.15 (0.29)	1.20 (0.17)	0.02	0.18 (0.09)	0.20 (0.07)	0.05	0.00 (0.00)	0.00 (0.00)	0.21
Quantile	0.11 (0.16)	0.00 (0.15)	0.76	0.01 (0.03)	0.00 (0.00)	2.41	0.00 (0.00)	0.00 (0.00)	10.41
EQC	0.00 (0.00)	0.00 (0.00)	0.23	0.00 (0.00)	0.00 (0.00)	0.79	0.00 (0.00)	0.00 (0.00)	3.48
nBayes	0.95 (0.26)	0.90 (0.22)	0.10	0.42 (0.19)	0.45 (0.22)	0.40	0.12 (0.11)	0.10 (0.15)	2.13
LDA	0.21 (0.12)	0.20 (0.15)	0.01	0.00 (0.00)	0.00 (0.00)	0.07	0.00 (0.00)	0.00 (0.00)	0.61
1-NN	20.65 (1.21)	20.45 (1.08)	0.01	18.40 (1.67)	18.20 (1.81)	0.01	14.38 (2.06)	14.60 (1.64)	0.05
SVM	1.45 (0.32)	1.40 (0.30)	0.01	0.13 (0.09)	0.10 (0.07)	0.03	0.00 (0.00)	0.00 (0.00)	0.20
AdaBoost	0.00 (0.00)	0.00 (0.00)	4.56	0.00 (0.00)	0.00 (0.00)	9.71	0.00 (0.00)	0.00 (0.00)	42.10
GBDT	0.00 (0.00)	0.00 (0.00)	4.91	0.00 (0.00)	0.00 (0.00)	7.93	0.00 (0.00)	0.00 (0.00)	36.24
XGBoost	3.09 (0.55)	3.00 (0.62)	0.01	2.84 (0.55)	2.80 (0.47)	0.02	2.97 (0.46)	2.90 (0.52)	0.03

Table S12: Misclassification rates multiplied by 100 for Case 5 of Example 1 with imbalanced setting and $\rho = 0.5$

	$p = 50$			$p = 200$			$p = 1000$		
	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time
$n = 50$ (imbalanced)									
Unimodal(ROT)	0.72 (0.59)	0.67 (1.00)	0.99	0.00 (0.00)	0.00 (0.00)	2.00	0.00 (0.00)	0.00 (0.00)	8.99
Unimodal(NS)	1.80 (0.86)	2.00 (0.62)	1.01	0.00 (0.00)	0.00 (0.00)	2.03	0.00 (0.00)	0.00 (0.00)	8.86
Quantile-mode	0.17 (0.29)	0.00 (0.12)	3.71	0.00 (0.00)	0.00 (0.00)	6.16	0.00 (0.00)	0.00 (0.00)	11.44
Naive-mode	30.72 (2.35)	31.33 (1.74)	0.07	33.13 (0.41)	33.33 (0.00)	0.10	33.33 (0.00)	33.33 (0.00)	0.22
Centroid	11.23 (3.04)	10.67 (3.48)	0.01	5.43 (1.60)	5.33 (1.62)	0.01	0.68 (0.63)	0.67 (1.00)	0.01
Median	1.05 (0.66)	0.67 (0.50)	0.01	0.38 (0.33)	0.67 (0.50)	0.06	0.00 (0.00)	0.00 (0.00)	0.20
Quantile	0.23 (0.39)	0.00 (0.50)	0.62	0.08 (0.22)	0.00 (0.00)	2.19	0.00 (0.00)	0.00 (0.00)	9.77
EQC	0.65 (0.11)	0.67 (0.00)	0.23	0.00 (0.00)	0.00 (0.00)	0.70	0.00 (0.00)	0.00 (0.00)	3.30
SVM	10.48 (3.56)	10.00 (2.99)	0.01	0.58 (0.63)	0.67 (0.50)	0.02	0.00 (0.00)	0.00 (0.00)	0.12
LDA	3.22 (1.98)	2.67 (1.49)	0.01	1.87 (1.39)	1.33 (1.12)	0.04	0.00 (0.00)	0.00 (0.00)	0.30
1-NN	23.05 (2.90)	23.33 (3.11)	0.01	18.82 (3.05)	18.33 (3.48)	0.01	14.15 (4.27)	13.67 (3.98)	0.01
nBayes	3.28 (1.38)	2.67 (1.12)	0.04	2.92 (1.41)	2.67 (1.12)	0.19	1.95 (0.80)	2.00 (1.00)	0.86
AdaBoost	0.25 (0.42)	0.00 (0.50)	4.05	0.05 (0.18)	0.00 (0.00)	7.70	0.00 (0.00)	0.00 (0.00)	32.52
GBDT	0.00 (0.00)	0.00 (0.00)	3.49	0.00 (0.00)	0.00 (0.00)	4.37	0.00 (0.00)	0.00 (0.00)	19.19
XGBoost	8.72 (2.63)	8.67 (2.24)	0.02	7.28 (3.18)	7.33 (2.49)	0.02	8.08 (2.69)	7.67 (2.49)	0.03
$n = 100$ (imbalanced)									
Unimodal(ROT)	0.94 (0.35)	0.83 (0.37)	1.13	0.03 (0.06)	0.00 (0.00)	2.06	0.00 (0.00)	0.00 (0.00)	10.01
Unimodal(NS)	3.07 (0.55)	3.08 (0.50)	0.87	0.38 (0.25)	0.33 (0.25)	2.25	0.00 (0.00)	0.00 (0.00)	9.37
Quantile-mode	0.07 (0.09)	0.00 (0.12)	4.36	0.00 (0.00)	0.00 (0.00)	7.48	0.00 (0.00)	0.00 (0.00)	14.22
Naive-mode	33.23 (0.97)	33.17 (0.12)	0.07	33.33 (0.00)	33.33 (0.00)	0.11	33.33 (0.00)	33.33 (0.00)	0.32
Centroid	9.97 (1.56)	10.00 (10.60)	0.01	3.73 (0.70)	3.67 (0.50)	0.01	0.45 (0.20)	0.50 (0.25)	0.01
Median	0.72 (0.30)	0.67 (0.25)	0.01	0.03 (0.07)	0.00 (0.00)	0.04	0.00 (0.00)	0.00 (0.00)	0.20
Quantile	0.08 (0.17)	0.00 (0.12)	0.71	0.00 (0.03)	0.00 (0.00)	2.37	0.00 (0.00)	0.00 (0.00)	10.14
EQC	0.00 (0.00)	0.00 (0.00)	0.24	0.00 (0.00)	0.00 (0.00)	0.76	0.00 (0.00)	0.00 (0.00)	3.47
SVM	1.88 (0.39)	1.83 (0.37)	0.01	0.23 (0.15)	0.25 (0.12)	0.03	0.00 (0.00)	0.00 (0.00)	0.17
LDA	0.13 (0.16)	0.17 (0.12)	0.02	0.01 (0.04)	0.00 (0.00)	0.06	0.23 (0.19)	0.25 (0.25)	0.49
1-NN	19.30 (1.06)	19.17 (1.15)	0.01	16.94 (1.26)	16.92 (1.06)	0.01	11.38 (1.23)	11.67 (1.52)	0.03
nBayes	1.30 (0.49)	1.17 (0.44)	0.08	0.73 (0.42)	0.67 (0.28)	0.31	0.35 (0.19)	0.33 (0.25)	1.67
AdaBoost	0.00 (0.03)	0.00 (0.00)	4.39	0.00 (0.00)	0.00 (0.00)	9.42	0.00 (0.00)	0.00 (0.00)	38.46
GBDT	0.00 (0.00)	0.00 (0.00)	4.21	0.00 (0.00)	0.00 (0.00)	6.36	0.00 (0.00)	0.00 (0.00)	28.66
XGBoost	2.60 (1.10)	2.33 (1.27)	0.01	2.93 (1.32)	2.58 (1.46)	0.02	4.25 (1.14)	4.50 (1.12)	0.03
$n = 200$ (imbalanced)									
Unimodal(ROT)	0.55 (0.16)	0.53 (0.10)	1.02	0.02 (0.04)	0.00 (0.05)	2.70	0.00 (0.00)	0.00 (0.00)	11.33
Unimodal(NS)	2.43 (0.37)	2.40 (0.35)	0.94	0.41 (0.12)	0.40 (0.11)	2.28	0.00 (0.00)	0.00 (0.00)	10.59
Quantile-mode	0.55 (0.12)	0.60 (0.15)	5.72	0.03 (0.03)	0.03 (0.05)	9.95	0.00 (0.00)	0.00 (0.00)	18.95
Naive-mode	33.31 (0.04)	33.33 (0.05)	0.08	33.33 (0.02)	33.33 (0.00)	0.15	33.33 (0.00)	33.33 (0.00)	0.53
Centroid	9.44 (0.81)	9.33 (0.76)	0.01	3.91 (0.52)	3.90 (0.42)	0.01	0.33 (0.12)	0.30 (0.11)	0.02
Median	0.93 (0.18)	0.93 (0.20)	0.02	0.10 (0.06)	0.07 (0.05)	0.05	0.00 (0.00)	0.00 (0.00)	0.23
Quantile	0.04 (0.05)	0.00 (0.05)	0.88	0.01 (0.03)	0.00 (0.00)	2.46	0.00 (0.00)	0.00 (0.00)	11.11
EQC	0.14 (0.02)	0.13 (0.00)	0.27	0.00 (0.00)	0.00 (0.00)	0.91	0.00 (0.00)	0.00 (0.00)	4.14
SVM	1.05 (0.17)	1.07 (0.25)	0.02	0.07 (0.06)	0.07 (0.06)	0.06	0.00 (0.00)	0.00 (0.00)	0.27
LDA	0.09 (0.07)	0.07 (0.10)	0.02	0.00 (0.00)	0.00 (0.00)	0.09	0.00 (0.00)	0.00 (0.00)	0.96
1-NN	18.55 (0.90)	18.70 (0.90)	0.01	16.92 (1.24)	16.83 (1.09)	0.02	11.98 (1.27)	11.97 (1.01)	0.10
nBayes	0.79 (0.22)	0.80 (0.16)	0.17	0.39 (0.15)	0.40 (0.15)	0.62	0.07 (0.08)	0.07 (0.10)	3.07
AdaBoost	0.00 (0.00)	0.00 (0.00)	5.00	0.00 (0.00)	0.00 (0.00)	11.62	0.00 (0.00)	0.00 (0.00)	52.16
GBDT	0.00 (0.00)	0.00 (0.00)	5.75	0.00 (0.00)	0.00 (0.00)	19.87	0.00 (0.00)	0.00 (0.00)	49.51
XGBoost	1.96 (0.40)	1.97 (0.31)	0.02	2.01 (0.32)	1.97 (0.35)	0.02	1.84 (0.36)	1.87 (0.30)	0.04

Table S13: Misclassification rates multiplied by 100 for Case 5 of Example 1 with balanced setting and $\rho = 0.8$

	$p = 50$			$p = 200$			$p = 1000$		
	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time
$n = 50$ (balanced)									
Unimodal(ROT)	2.33 (1.31)	2.00 (0.93)	1.02	0.08 (0.27)	0.00 (0.00)	2.18	0.00 (0.00)	0.00 (0.00)	8.92
Unimodal(NS)	4.20 (1.60)	4.00 (1.49)	0.87	0.63 (0.63)	1.00 (0.75)	1.83	0.00 (0.00)	0.00 (0.00)	9.13
Quantile-mode	0.83 (0.71)	1.00 (0.75)	3.53	0.00 (0.00)	0.00 (0.00)	5.66	0.00 (0.00)	0.00 (0.00)	10.97
Naive-mode	44.38 (5.20)	46.00 (5.97)	0.06	42.65 (4.62)	43.50 (4.85)	0.10	49.73 (0.60)	50.00 (50.00)	0.17
Centroid	22.20 (3.96)	21.50 (4.48)	0.01	14.60 (3.25)	15.00 (2.43)	0.01	6.40 (2.73)	6.00 (2.43)	0.01
Median	8.70 (2.76)	8.00 (2.99)	0.01	2.48 (1.30)	2.00 (0.75)	0.05	0.00 (0.00)	0.00 (0.00)	0.20
Quantile	2.05 (1.91)	1.50 (1.49)	0.60	0.30 (0.52)	0.00 (0.75)	2.03	0.03 (0.16)	0.00 (0.00)	9.72
EQC	1.98 (0.16)	2.00 (0.00)	0.21	0.00 (0.00)	0.00 (0.00)	0.63	0.00 (0.00)	0.00 (0.00)	3.11
SVM	12.65 (2.96)	13.00 (2.24)	0.01	4.30 (1.52)	4.00 (0.75)	0.02	0.03 (0.16)	0.00 (0.00)	0.11
LDA	37.43 (8.29)	37.50 (8.21)	0.01	2.33 (1.59)	2.00 (1.68)	0.03	0.10 (0.30)	0.00 (0.00)	0.25
1-NN	27.10 (4.59)	27.00 (3.73)	0.01	24.10 (3.72)	24.50 (2.99)	0.01	31.30 (5.71)	32.50 (4.85)	0.01
nBayes	9.60 (2.92)	10.00 (3.17)	0.04	4.53 (2.20)	4.00 (2.24)	0.14	3.50 (1.60)	3.00 (1.49)	0.63
AdaBoost	1.35 (1.19)	1.00 (0.93)	4.00	0.30 (0.61)	0.00 (0.00)	7.54	0.10 (0.30)	0.00 (0.00)	31.47
GBDT	0.25 (0.54)	0.00 (0.00)	3.59	0.15 (0.36)	0.00 (0.00)	4.19	0.03 (0.16)	0.00 (0.00)	18.61
XGBoost	15.03 (4.18)	15.00 (2.99)	0.01	16.38 (4.18)	16.50 (2.99)	0.02	16.23 (3.66)	16.50 (3.36)	0.02
$n = 100$ (balanced)									
Unimodal(ROT)	0.80 (0.37)	0.75 (0.37)	0.83	0.00 (0.00)	0.00 (0.00)	2.14	0.00 (0.00)	0.00 (0.00)	9.48
Unimodal(NS)	2.60 (0.66)	2.50 (0.79)	0.94	0.15 (0.16)	0.25 (0.19)	2.01	0.00 (0.00)	0.00 (0.00)	9.02
Quantile-mode	0.56 (0.29)	0.50 (0.23)	3.70	0.00 (0.00)	0.00 (0.00)	6.54	0.00 (0.00)	0.00 (0.00)	12.30
Naive-mode	49.84 (0.23)	50.00 (0.19)	0.06	50.00 (0.00)	50.00 (0.00)	0.10	50.00 (0.00)	50.00 (0.00)	0.24
Centroid	9.61 (1.40)	9.75 (1.49)	0.01	4.40 (0.87)	4.50 (0.75)	0.01	0.61 (0.33)	0.50 (0.23)	0.01
Median	1.75 (0.52)	1.75 (0.37)	0.01	0.02 (0.07)	0.00 (0.00)	0.04	0.00 (0.00)	0.00 (0.00)	0.19
Quantile	0.20 (0.21)	0.25 (0.19)	0.66	0.05 (0.20)	0.00 (0.00)	2.15	0.00 (0.00)	0.00 (0.00)	9.63
EQC	0.26 (0.04)	0.25 (0.00)	0.23	0.00 (0.00)	0.00 (0.00)	0.74	0.00 (0.00)	0.00 (0.00)	3.12
nBayes	1.88 (0.66)	1.75 (0.75)	0.07	0.64 (0.41)	0.50 (0.51)	0.23	0.59 (0.31)	0.50 (0.19)	1.19
LDA	0.27 (0.21)	0.25 (0.37)	0.01	0.18 (0.23)	0.13 (0.19)	0.05	0.22 (0.27)	0.25 (0.23)	0.36
1-NN	21.21 (1.68)	20.88 (1.91)	0.01	20.61 (1.33)	20.38 (1.17)	0.01	16.73 (2.54)	17.50 (2.71)	0.01
SVM	2.50 (0.59)	2.50 (0.61)	0.01	0.23 (0.18)	0.25 (0.14)	0.04	0.00 (0.00)	0.00 (0.00)	0.13
AdaBoost	0.00 (0.00)	0.00 (0.00)	4.22	0.00 (0.00)	0.00 (0.00)	8.46	0.00 (0.00)	0.00 (0.00)	34.59
GBDT	0.00 (0.00)	0.00 (0.00)	4.20	0.00 (0.00)	0.00 (0.00)	4.64	0.00 (0.00)	0.00 (0.00)	22.73
XGBoost	3.70 (1.86)	3.25 (1.40)	0.23	3.62 (1.73)	3.00 (1.87)	0.74	4.79 (2.72)	1.41 (2.24)	3.12
$n = 200$ (balanced)									
Unimodal(ROT)	2.75 (0.46)	2.75 (0.47)	1.04	0.29 (0.12)	0.30 (0.15)	2.68	0.00 (0.00)	0.00 (0.00)	10.43
Unimodal(NS)	6.85 (0.87)	6.90 (0.97)	0.99	1.46 (0.31)	1.40 (0.24)	1.71	0.00 (0.00)	0.00 (0.00)	9.68
Quantile-mode	2.91 (0.46)	2.90 (0.14)	4.70	0.06 (0.06)	0.10 (0.07)	8.21	0.00 (0.00)	0.00 (0.00)	15.55
Naive-mode	48.96 (0.42)	48.95 (0.47)	0.08	49.63 (0.31)	49.80 (0.30)	0.14	50.00 (0.00)	50.00 (0.00)	0.42
Centroid	18.47 (1.20)	1.83 (1.32)	0.01	13.15 (1.29)	13.20 (1.08)	0.01	4.86 (0.73)	4.80 (0.82)	0.01
Median	6.30 (0.74)	6.30 (0.86)	0.02	1.85 (0.36)	1.90 (0.34)	0.05	0.16 (0.09)	0.20 (0.07)	0.25
Quantile	0.77 (0.29)	0.80 (0.32)	0.78	0.07 (0.12)	0.00 (0.07)	2.45	0.02 (0.04)	0.00 (0.00)	10.75
EQC	0.49 (0.08)	0.50 (0.00)	0.28	0.10 (0.00)	0.10 (0.00)	0.76	0.00 (0.00)	0.00 (0.00)	3.66
SVM	4.75 (0.68)	4.60 (0.71)	0.02	0.67 (0.24)	0.60 (0.22)	0.04	0.00 (0.00)	0.00 (0.00)	0.19
LDA	3.04 (0.47)	3.10 (0.37)	0.02	0.50 (0.21)	0.50 (0.19)	0.08	0.16 (0.15)	0.10 (0.17)	0.61
1-NN	22.74 (1.31)	22.50 (1.23)	0.01	22.13 (1.46)	22.00 (1.57)	0.01	18.53 (1.37)	18.60 (1.23)	0.05
nBayes	4.19 (0.75)	4.20 (0.69)	0.12	1.08 (0.45)	1.00 (0.52)	0.41	0.27 (0.15)	0.30 (0.15)	2.08
AdaBoost	0.05 (0.07)	0.00 (0.07)	4.52	0.00 (0.02)	0.00 (0.00)	9.63	0.00 (0.00)	0.00 (0.00)	42.42
GBDT	0.00 (0.00)	0.00 (0.00)	5.02	0.00 (0.00)	0.00 (0.00)	7.82	0.00 (0.00)	0.00 (0.00)	36.48
XGBoost	2.93 (0.65)	2.85 (0.67)	0.01	2.93 (0.61)	2.80 (0.63)	0.02	2.76 (0.45)	2.75 (0.52)	0.03

Table S14: Misclassification rates multiplied by 100 for Case 5 of Example 1 with imbalanced setting and $\rho = 0.8$

	$p = 50$			$p = 200$			$p = 1000$		
	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time
$n = 50$ (imbalanced)									
Unimodal(ROT)	0.72 (0.73)	0.67 (1.00)	1.09	0.10 (0.24)	0.00 (0.00)	2.10	0.00 (0.00)	0.00 (0.00)	9.04
Unimodal(NS)	2.45 (1.18)	2.00 (1.12)	1.08	0.52 (0.53)	0.67 (0.50)	2.18	0.00 (0.00)	0.00 (0.00)	8.47
Quantile-mode	0.58 (0.48)	0.67 (0.50)	3.78	0.03 (0.15)	0.00 (0.00)	6.25	0.00 (0.00)	0.00 (0.00)	11.85
Naive-mode	31.75 (1.36)	32.00 (1.00)	0.07	32.87 (0.43)	32.67 (0.50)	0.10	33.27 (0.20)	33.33 (0.00)	0.23
Centroid	18.60 (2.58)	18.00 (2.11)	0.01	13.78 (2.83)	13.67 (2.99)	0.01	5.43 (1.76)	5.33 (1.62)	0.01
Median	5.02 (1.48)	5.00 (1.49)	0.01	2.60 (1.20)	2.67 (1.00)	0.04	0.32 (0.34)	0.00 (0.50)	0.21
Quantile	0.82 (1.02)	0.67 (1.00)	0.62	0.55 (0.56)	0.67 (0.50)	2.19	0.00 (0.00)	0.00 (0.00)	9.64
EQC	0.03 (0.21)	0.00 (0.00)	0.20	0.02 (0.11)	0.00 (0.00)	0.67	0.00 (0.00)	0.00 (0.00)	3.12
SVM	14.77 (3.79)	14.67 (4.48)	0.01	2.93 (1.22)	3.33 (1.49)	0.02	0.03 (0.15)	0.00 (0.00)	0.14
LDA	12.95 (3.34)	12.67 (2.99)	0.01	7.20 (2.50)	7.33 (1.62)	0.05	0.05 (0.18)	0.00 (0.00)	0.30
1-NN	18.35 (2.53)	18.67 (2.11)	0.01	22.22 (2.62)	22.00 (2.49)	0.01	19.23 (2.66)	18.67 (3.11)	0.01
nBayes	4.70 (1.87)	4.67 (1.62)	0.05	3.58 (1.68)	3.33 (1.49)	0.17	1.30 (1.06)	1.33 (1.00)	0.89
AdaBoost	0.91 (1.36)	0.67 (0.50)	3.83	0.13 (0.31)	0.00 (0.00)	7.61	0.00 (0.00)	0.00 (0.00)	33.05
GBDT	0.00 (0.00)	0.00 (0.00)	4.44	0.00 (0.00)	0.00 (0.00)	4.57	0.00 (0.00)	0.00 (0.00)	18.21
XGBoost	8.70 (2.33)	8.67 (1.49)	0.20	8.90 (2.72)	8.67 (2.11)	0.67	9.13 (1.90)	9.00 (2.11)	3.12
$n = 100$ (imbalanced)									
Unimodal(ROT)	1.64 (0.47)	1.67 (0.37)	0.87	0.14 (0.12)	0.17 (0.12)	1.72	0.00 (0.00)	0.00 (0.00)	9.97
Unimodal(NS)	4.60 (0.66)	4.58 (0.75)	0.96	1.35 (0.38)	1.33 (0.28)	2.17	0.01 (0.04)	0.00 (0.00)	9.19
Quantile-mode	1.42 (0.37)	1.50 (0.37)	4.41	0.05 (0.08)	0.00 (0.12)	7.40	0.00 (0.00)	0.00 (0.00)	14.06
Naive-mode	32.26 (0.71)	32.33 (0.56)	0.08	33.20 (0.19)	33.33 (0.16)	0.13	33.33 (0.00)	33.33 (0.00)	0.32
Centroid	19.40 (2.12)	19.08 (2.02)	0.01	13.51 (1.56)	13.58 (1.40)	0.01	5.36 (0.80)	5.33 (0.68)	0.01
Median	5.05 (0.92)	5.08 (0.78)	0.01	1.65 (0.57)	1.67 (0.40)	0.06	0.30 (0.17)	0.33 (0.25)	0.20
Quantile	0.48 (0.39)	0.42 (0.37)	0.70	0.12 (0.15)	0.00 (0.12)	2.30	0.00 (0.03)	0.00 (0.00)	10.20
EQC	0.18 (0.08)	0.17 (0.00)	0.24	0.02 (0.11)	0.00 (0.00)	0.84	0.00 (0.00)	0.00 (0.00)	3.55
SVM	6.30 (0.69)	6.25 (0.62)	0.01	1.32 (0.31)	1.25 (0.25)	0.03	0.00 (0.03)	0.00 (0.00)	0.17
LDA	2.90 (0.59)	2.83 (0.62)	0.02	1.24 (0.37)	1.17 (0.37)	0.06	3.81 (1.41)	3.67 (1.43)	0.49
1-NN	19.55 (0.96)	19.58 (1.00)	0.01	19.83 (1.46)	19.67 (1.74)	0.01	18.48 (1.76)	18.33 (1.90)	0.03
nBayes	3.32 (0.86)	3.17 (0.90)	0.08	1.53 (0.58)	1.50 (0.53)	0.31	0.40 (0.29)	0.33 (0.25)	1.64
AdaBoost	0.03 (0.07)	0.00 (0.00)	4.32	0.03 (0.07)	0.00 (0.00)	9.56	0.00 (0.00)	0.00 (0.00)	38.42
GBDT	0.00 (0.00)	0.00 (0.00)	4.60	0.00 (0.00)	0.00 (0.00)	6.50	0.00 (0.00)	0.00 (0.00)	28.29
XGBoost	2.58 (1.13)	2.25 (1.15)	0.01	2.93 (1.00)	2.67 (1.06)	0.02	3.73 (1.14)	3.67 (1.06)	0.03
$n = 200$ (imbalanced)									
Unimodal(ROT)	1.82 (0.25)	1.80 (0.21)	1.05	0.33 (0.10)	0.33 (0.10)	2.15	0.00 (0.00)	0.00 (0.00)	12.00
Unimodal(NS)	4.45 (0.51)	4.47 (0.41)	1.07	1.53 (0.26)	1.53 (0.25)	2.08	0.03 (0.03)	0.00 (0.05)	10.65
Quantile-mode	2.11 (0.30)	2.13 (0.22)	5.74	0.25 (0.08)	0.27 (0.10)	9.93	0.00 (0.00)	0.00 (0.00)	18.85
Naive-mode	31.55 (0.91)	31.60 (0.80)	0.09	33.16 (0.18)	33.20 (0.10)	0.17	33.33 (0.00)	33.33 (0.00)	0.52
Centroid	20.40 (1.10)	20.30 (1.06)	0.01	13.19 (0.69)	13.17 (0.66)	0.01	4.40 (0.55)	4.33 (0.72)	0.02
Median	5.03 (0.48)	4.93 (0.57)	0.02	1.89 (0.23)	1.87 (0.19)	0.07	0.01 (0.00)	0.00 (0.00)	0.22
Quantile	0.49 (0.17)	0.47 (0.47)	0.86	0.04 (0.06)	0.00 (0.05)	2.47	0.00 (0.00)	0.00 (0.00)	11.17
EQC	0.27 (0.00)	0.27 (0.00)	0.33	0.00 (0.00)	0.00 (0.00)	0.95	0.00 (0.00)	0.00 (0.00)	4.32
SVM	3.18 (0.23)	3.20 (0.16)	0.03	0.76 (0.17)	0.73 (0.17)	0.06	0.00 (0.00)	0.00 (0.00)	0.25
LDA	1.76 (0.26)	1.80 (0.26)	0.02	0.40 (0.16)	0.40 (0.11)	0.09	0.02 (0.03)	0.00 (0.01)	0.96
1-NN	19.04 (0.83)	19.13 (0.87)	0.01	19.30 (0.90)	19.20 (0.63)	0.02	15.87 (0.91)	15.70 (0.95)	0.11
nBayes	2.73 (0.54)	2.80 (0.47)	0.17	1.10 (0.25)	1.07 (0.30)	0.60	0.14 (0.10)	0.13 (0.10)	3.01
AdaBoost	0.02 (0.03)	0.00 (0.01)	5.15	0.03 (0.03)	0.00 (0.05)	11.85	0.00 (0.00)	0.00 (0.00)	51.65
GBDT	0.00 (0.00)	0.00 (0.00)	5.81	0.00 (0.00)	0.00 (0.00)	20.32	0.00 (0.00)	0.00 (0.00)	51.01
XGBoost	2.25 (0.47)	2.27 (0.57)	0.01	2.01 (0.44)	2.00 (0.57)	0.02	1.95 (0.31)	1.93 (0.31)	0.04

Table S15: Misclassification rates multiplied by 100 for Case 6 of Example 1 with balanced setting

	p = 50			p = 200			p = 1000		
	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time
$n = 50$ (balanced)									
Unimodal(ROT)	2.55 (2.55)	1.20 (3.00)	1.45	0.20 (0.46)	0.00 (0.00)	2.93	0.00 (0.00)	0.00 (0.00)	11.71
Unimodal(NS)	1.65 (1.21)	2.00 (0.75)	1.27	0.15 (0.36)	0.00 (0.00)	2.47	0.00 (0.00)	0.00 (0.00)	11.10
Quantile-mode	0.18 (0.38)	0.00 (0.00)	2.95	0.00 (0.00)	0.00 (0.00)	5.53	0.00 (0.00)	0.00 (0.00)	9.76
Naive-mode	30.58 (4.95)	31.00 (5.41)	0.06	34.38 (4.02)	35.00 (4.66)	0.08	40.53 (3.41)	41.00 (2.99)	0.14
Centroid	8.63 (1.94)	9.00 (2.24)	0.01	0.05 (0.22)	0.00 (0.00)	0.01	0.00 (0.00)	0.00 (0.00)	0.01
Median	6.88 (1.81)	7.00 (1.45)	0.01	0.03 (0.16)	0.00 (0.00)	0.04	0.00 (0.00)	0.00 (0.00)	0.19
Quantile	2.50 (2.30)	2.00 (2.24)	0.62	1.25 (1.74)	1.00 (1.49)	2.11	0.08 (0.27)	0.00 (0.00)	9.81
EQC	6.90 (0.63)	7.00 (0.00)	0.21	0.00 (0.00)	0.00 (0.00)	0.65	0.00 (0.00)	0.00 (0.00)	2.88
SVM	4.08 (1.47)	4.00 (1.49)	0.02	0.00 (0.00)	0.00 (0.00)	0.03	0.00 (0.00)	0.00 (0.00)	0.12
LDA	16.00 (3.74)	16.50 (2.99)	0.01	8.68 (3.16)	8.50 (3.36)	0.04	5.43 (2.58)	5.50 (2.99)	0.31
1-NN	27.25 (5.07)	27.00 (5.41)	0.01	11.50 (4.06)	10.50 (3.92)	0.01	2.90 (2.38)	2.00 (2.99)	0.01
nBayes	19.20 (3.41)	20.00 (3.73)	0.04	6.28 (2.03)	6.00 (2.24)	0.14	1.33 (1.12)	1.00 (0.75)	0.71
AdaBoost	3.60 (1.92)	3.00 (1.68)	4.00	0.93 (1.10)	1.00 (0.75)	7.78	1.00 (1.11)	1.00 (1.49)	33.14
GBDT	3.98 (1.66)	4.00 (1.49)	3.41	0.25 (0.54)	0.00 (0.00)	4.05	0.03 (0.16)	0.00 (0.00)	17.95
XGBoost	22.03 (5.49)	22.00 (4.66)	0.01	24.03 (5.13)	24.50 (5.22)	0.02	22.95 (4.28)	23.00 (4.66)	0.02
$n = 100$ (balanced)									
Unimodal(ROT)	6.49 (1.04)	6.38 (1.12)	1.71	0.08 (0.12)	0.00 (0.19)	3.02	0.00 (0.00)	0.00 (0.00)	13.98
Unimodal(NS)	5.77 (0.97)	5.75 (0.56)	1.50	0.04 (0.11)	0.00 (0.00)	2.31	0.00 (0.00)	0.00 (0.00)	11.24
Quantile-mode	0.78 (0.34)	0.75 (0.37)	3.23	0.00 (0.00)	0.00 (0.00)	5.57	0.00 (0.00)	0.00 (0.00)	10.00
Naive-mode	26.84 (2.38)	26.75 (2.71)	0.07	32.92 (2.77)	33.13 (2.66)	0.08	37.44 (2.28)	37.88 (2.15)	0.16
Centroid	7.21 (0.92)	7.50 (0.93)	0.01	0.13 (0.15)	0.00 (0.19)	0.01	0.00 (0.00)	0.00 (0.00)	0.01
Median	7.31 (1.10)	7.25 (1.31)	0.02	0.27 (0.18)	0.25 (0.00)	0.05	0.00 (0.00)	0.00 (0.00)	0.23
Quantile	0.03 (0.08)	0.00 (0.00)	0.72	0.00 (0.00)	0.00 (0.00)	2.13	0.00 (0.00)	0.00 (0.00)	10.11
EQC	12.90 (0.24)	7.00 (0.00)	0.23	0.00 (0.00)	0.00 (0.00)	0.73	0.00 (0.00)	0.00 (0.00)	3.28
SVM	5.94 (0.98)	6.00 (1.17)	0.01	0.00 (0.00)	0.00 (0.00)	0.03	0.00 (0.00)	0.00 (0.00)	0.17
LDA	9.11 (1.25)	9.00 (1.12)	0.01	1.16 (0.62)	7.50 (0.37)	0.05	2.29 (0.95)	2.25 (0.93)	0.38
1-NN	22.89 (2.23)	23.25 (2.29)	0.01	10.61 (2.72)	10.50 (0.27)	0.01	1.91 (0.81)	2.00 (0.93)	0.02
nBayes	11.14 (1.53)	11.38 (1.73)	0.08	1.22 (0.58)	1.25 (0.75)	0.25	0.02 (0.07)	0.00 (0.00)	1.18
AdaBoost	0.64 (0.40)	0.50 (0.42)	4.54	0.01 (0.04)	0.00 (0.00)	8.42	0.00 (0.00)	0.00 (0.00)	35.89
GBDT	0.38 (0.37)	0.25 (0.23)	4.98	0.00 (0.00)	0.00 (0.00)	4..71	0.00 (0.00)	0.00 (0.00)	23.49
XGBoost	14.17 (2.23)	13.75 (1.87)	0.01	14.76 (2.94)	14.25 (3.08)	0.02	14.31 (2.13)	14.63 (2.47)	0.02
$n = 200$ (balanced)									
Unimodal(ROT)	6.77 (0.74)	6.75 (0.75)	1.54	0.19 (0.12)	0.20 (0.15)	3.64	0.00 (0.00)	0.00 (0.00)	13.79
Unimodal(NS)	6.39 (0.70)	6.45 (0.82)	1.69	0.14 (0.10)	0.10 (0.07)	2.59	0.00 (0.00)	0.00 (0.00)	12.09
Quantile-mode	1.45 (0.25)	1.45 (0.24)	3.53	0.00 (0.00)	0.00 (0.00)	5.79	0.00 (0.00)	0.00 (0.00)	11.03
Naïve-mode	28.14 (1.29)	28.05 (1.38)	0.07	31.39 (1.82)	31.45 (1.88)	0.09	36.57 (1.49)	36.40 (1.72)	0.20
Centroid	6.92 (0.50)	6.90 (0.54)	0.01	0.15 (0.07)	0.20 (0.07)	0.01	0.00 (0.00)	0.00 (0.00)	0.01
Median	7.76 (0.66)	7.80 (0.67)	0.02	0.15 (0.09)	0.10 (0.07)	0.07	0.00 (0.00)	0.00 (0.00)	0.24
Quantile	0.04 (0.05)	0.00 (0.07)	0.85	0.00 (0.00)	0.00 (0.00)	2.48	0.00 (0.00)	0.00 (0.00)	10.96
EQC	13.10 (0.05)	13.00 (0.00)	0.28	0.00 (0.00)	0.00 (0.00)	0.86	0.00 (0.00)	0.00 (0.00)	3.63
SVM	6.14 (0.61)	6.10 (0.63)	0.02	0.01 (0.03)	0.00 (0.00)	0.04	0.00 (0.00)	0.00 (0.00)	0.21
LDA	7.53 (0.52)	7.55 (0.56)	0.02	0.35 (0.17)	0.40 (0.19)	0.09	0.06 (0.05)	0.10 (0.07)	0.65
1-NN	22.59 (1.40)	22.55 (1.19)	0.01	6.41 (0.57)	6.35 (0.52)	0.01	2.00 (0.77)	1.90 (0.71)	0.05
nBayes	9.62 (1.12)	9.45 (1.12)	0.13	0.35 (0.17)	0.35 (0.15)	0.66	0.00 (0.00)	0.00 (0.00)	2.14
AdaBoost	0.17 (0.15)	0.20 (0.22)	5.46	0.00 (0.00)	0.00 (0.00)	11.39	0.00 (0.00)	0.00 (0.00)	47.06
GBDT	0.14 (0.12)	0.10 (0.15)	5.81	0.00 (0.00)	0.00 (0.00)	8.84	0.00 (0.00)	0.00 (0.00)	39.00
XGBoost	11.11 (1.28)	11.10 (1.06)	0.01	11.30 (1.32)	11.30 (1.64)	0.02	11.91 (0.79)	11.85 (0.69)	0.03

Table S16: Misclassification rates multiplied by 100 for Case 6 of Example 1 with imbalanced setting

	$p = 50$			$p = 200$			$p = 1000$		
	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time
$n = 50$ (imbalanced)									
Unimodal(ROT)	4.33 (1.54)	4.00 (1.12)	1.49	0.07 (0.20)	0.00 (0.00)	3.32	0.00 (0.00)	0.00 (0.00)	13.20
Unimodal(NS)	3.48 (1.10)	3.67 (1.00)	1.31	0.00 (0.00)	0.00 (0.00)	3.31	0.00 (0.00)	0.00 (0.00)	11.48
Quantile-mode	0.68 (0.44)	0.67 (0.00)	2.90	0.00 (0.00)	0.00 (0.00)	5.28	0.00 (0.00)	0.00 (0.00)	9.84
Naive-mode	24.68 (3.61)	24.67 (4.48)	0.06	28.83 (2.24)	29.00 (2.99)	0.08	32.32 (0.84)	32.67 (0.62)	0.14
Centroid	7.20 (1.79)	7.33 (1.49)	0.01	0.08 (0.22)	0.00 (0.00)	0.01	0.00 (0.00)	0.00 (0.00)	0.01
Median	6.77 (1.35)	6.67 (1.49)	0.02	0.15 (0.28)	0.00 (0.00)	0.06	0.00 (0.00)	0.00 (0.00)	0.23
Quantile	0.53 (0.92)	0.00 (0.50)	0.71	0.15 (0.44)	0.00 (0.00)	2.45	0.00 (0.00)	0.00 (0.00)	10.31
EQC	1.98 (0.11)	2.00 (0.00)	0.23	0.02 (0.11)	0.00 (0.00)	0.72	0.00 (0.00)	0.00 (0.00)	3.17
SVM	5.33 (1.27)	5.33 (1.00)	0.01	0.00 (0.00)	0.00 (0.00)	0.02	0.00 (0.00)	0.00 (0.00)	0.13
LDA	10.53 (2.15)	10.33 (2.49)	0.02	10.37 (3.71)	9.67 (4.10)	0.04	2.62 (1.36)	2.67 (1.12)	0.34
1-NN	19.50 (2.02)	19.33 (1.99)	0.01	15.50 (3.90)	15.33 (3.48)	0.01	5.88 (2.40)	5.33 (1.99)	0.01
nBayes	15.02 (2.73)	14.67 (2.49)	0.06	5.03 (1.67)	5.33 (1.99)	0.20	1.80 (1.06)	2.00 (1.49)	0.98
AdaBoost	3.35 (1.74)	3.33 (1.99)	4.10	1.00 (0.66)	1.33 (0.50)	8.04	0.55 (0.62)	0.67 (0.62)	33.16
GBDT	1.78 (1.51)	1.33 (1.00)	4.74	0.00 (0.00)	0.00 (0.00)	5.13	0.00 (0.00)	0.00 (0.00)	19.67
XGBoost	15.75 (3.67)	15.67 (3.98)	0.01	15.87 (2.92)	15.67 (3.11)	0.02	16.40 (3.34)	16.00 (2.49)	0.02
$n = 100$ (imbalanced)									
Unimodal(ROT)	4.43 (0.85)	4.17 (0.75)	1.51	0.04 (0.08)	0.00 (0.00)	3.26	0.00 (0.00)	0.00 (0.00)	14.22
Unimodal(NS)	4.18 (0.68)	4.00 (0.75)	1.21	0.01 (0.04)	0.00 (0.00)	2.94	0.00 (0.00)	0.00 (0.00)	11.79
Quantile-mode	0.53 (0.20)	0.50 (0.25)	3.46	0.00 (0.00)	0.00 (0.00)	5.77	0.00 (0.00)	0.00 (0.00)	10.57
Naive-mode	20.08 (1.77)	20.00 (1.43)	0.08	23.82 (1.20)	23.75 (0.93)	0.08	27.29 (0.94)	27.17 (0.78)	0.17
Centroid	6.70 (0.96)	6.58 (1.03)	0.01	0.12 (0.13)	0.17 (0.12)	0.01	0.00 (0.00)	0.00 (0.00)	0.01
Median	6.58 (0.72)	6.50 (0.62)	0.02	0.14 (0.14)	0.17 (0.16)	0.05	0.00 (0.00)	0.00 (0.00)	0.24
Quantile	0.00 (0.03)	0.00 (0.00)	0.79	0.00 (0.00)	0.00 (0.00)	2.61	0.00 (0.00)	0.00 (0.00)	10.70
EQC	1.34 (0.05)	1.33 (0.00)	0.26	0.16 (0.03)	0.17 (0.00)	0.82	0.00 (0.00)	0.00 (0.00)	3.61
SVM	4.65 (0.67)	4.50 (0.78)	0.01	0.00 (0.00)	0.00 (0.00)	0.03	0.00 (0.00)	0.00 (0.00)	0.18
LDA	6.88 (0.89)	6.67 (1.03)	0.01	0.87 (0.35)	0.83 (0.37)	0.06	2.22 (0.86)	2.00 (0.75)	0.49
1-NN	20.27 (1.35)	20.33 (1.31)	0.01	9.68 (1.34)	9.75 (1.55)	0.01	2.34 (0.95)	2.33 (1.24)	0.03
nBayes	9.63 (1.30)	9.42 (1.62)	0.10	1.03 (0.33)	1.00 (0.28)	0.35	0.00 (0.00)	0.00 (0.00)	1.63
AdaBoost	0.68 (0.32)	0.67 (0.25)	4.69	0.02 (0.05)	0.00 (0.00)	9.51	0.00 (0.00)	0.00 (0.00)	40.15
GBDT	0.34 (0.21)	0.33 (0.25)	5.07	0.00 (0.00)	0.00 (0.00)	6.87	0.00 (0.00)	0.00 (0.00)	29.89
XGBoost	8.60 (1.66)	8.50 (1.43)	0.01	8.69 (1.95)	8.25 (1.90)	0.02	8.80 (1.36)	8.75 (1.49)	0.03
$n = 200$ (imbalanced)									
Unimodal(ROT)	5.46 (0.47)	5.50 (0.60)	1.67	0.16 (0.08)	0.13 (0.06)	4.23	0.00 (0.00)	0.00 (0.00)	14.58
Unimodal(NS)	5.15 (0.47)	5.17 (0.55)	1.55	0.12 (0.06)	0.13 (0.05)	3.59	0.00 (0.00)	0.00 (0.00)	13.78
Quantile-mode	1.32 (0.26)	1.30 (0.25)	3.78	0.00 (0.00)	0.00 (0.00)	6.46	0.00 (0.00)	0.00 (0.00)	11.99
Naive-mode	19.87 (1.10)	20.00 (1.38)	0.07	21.95 (0.79)	21.87 (0.78)	0.09	25.23 (0.64)	25.23 (0.71)	0.23
Centroid	6.29 (0.50)	6.30 (0.52)	0.01	0.16 (0.09)	0.13 (0.11)	0.01	0.00 (0.00)	0.00 (0.00)	0.02
Median	6.24 (0.45)	6.27 (0.62)	0.02	0.15 (0.08)	0.13 (0.10)	0.05	0.00 (0.00)	0.00 (0.00)	0.27
Quantile	0.00 (0.00)	0.00 (0.00)	1.28	0.00 (0.00)	0.00 (0.00)	2.61	0.00 (0.00)	0.00 (0.00)	11.77
EQC	1.39 (0.04)	1.40 (0.00)	0.34	0.00 (0.00)	0.00 (0.00)	0.91	0.00 (0.00)	0.00 (0.00)	4.31
SVM	5.32 (0.51)	5.23 (0.55)	0.03	0.02 (0.04)	0.00 (0.05)	0.04	0.00 (0.00)	0.00 (0.00)	0.30
LDA	5.93 (0.53)	5.93 (0.52)	0.02	0.25 (0.10)	0.27 (0.10)	0.11	0.00 (0.01)	0.00 (0.00)	1.03
1-NN	18.35 (0.80)	18.40 (0.88)	0.01	6.08 (0.78)	6.07 (0.85)	0.02	1.70 (0.61)	1.53 (0.61)	0.12
nBayes	6.99 (0.67)	7.03 (0.73)	0.19	0.40 (0.12)	0.43 (0.10)	0.62	0.00 (0.01)	0.00 (0.00)	3.06
AdaBoost	0.11 (0.09)	0.07 (0.05)	5.69	0.00 (0.00)	0.00 (0.00)	12.95	0.00 (0.00)	0.00 (0.00)	58.49
GBDT	0.05 (0.04)	0.07 (0.05)	6.14	0.00 (0.00)	0.00 (0.00)	20.83	0.00 (0.00)	0.00 (0.00)	53.09
XGBoost	8.13 (0.84)	8.13 (0.85)	0.01	7.98 (0.80)	7.93 (0.65)	0.02	8.20 (0.74)	8.10 (0.86)	0.04

Table S17: Misclassification rates multiplied by 100 for Case 7 of Example 1 with balanced setting

	$p = 50$			$p = 200$			$p = 1000$		
	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time
$n = 50$ (balanced)									
Unimodal(ROT)	2.75 (1.46)	3.00 (1.12)	1.02	0.00 (0.00)	0.00 (0.00)	2.36	0.00 (0.00)	0.00 (0.00)	10.50
Unimodal(NS)	2.70 (1.44)	3.00 (0.93)	0.77	0.00 (0.00)	0.00 (0.00)	2.10	0.00 (0.00)	0.00 (0.00)	9.19
Quantile-mode	0.25 (0.44)	0.00 (0.19)	2.93	0.00 (0.00)	0.00 (0.00)	5.16	0.00 (0.00)	0.00 (0.00)	9.56
Naive-mode	18.45 (4.55)	18.50 (3.36)	0.07	11.58 (3.13)	11.50 (2.43)	0.07	5.53 (2.50)	5.50 (2.24)	0.13
Centroid	41.70 (3.39)	41.00 (2.43)	0.01	40.83 (2.75)	41.00 (1.68)	0.01	44.70 (2.10)	44.50 (1.68)	0.01
Median	17.00 (3.06)	17.00 (2.99)	0.02	8.48 (2.44)	8.00 (2.24)	0.05	1.85 (1.27)	2.00 (0.93)	0.23
Quantile	4.18 (2.85)	3.50 (2.99)	0.68	1.35 (1.19)	1.00 (1.49)	2.65	0.05 (0.22)	0.00 (0.00)	9.77
EQC	1.03 (0.16)	1.00 (0.00)	0.24	0.00 (0.00)	0.00 (0.00)	0.68	0.00 (0.00)	0.00 (0.00)	3.01
SVM	31.38 (5.70)	31.00 (5.41)	0.01	9.93 (2.56)	10.00 (2.43)	0.02	0.08 (0.27)	0.00 (0.00)	0.13
LDA	42.68 (4.26)	43.00 (2.61)	0.01	41.45 (4.49)	41.00 (4.10)	0.04	42.60 (3.46)	42.00 (3.92)	0.31
1-NN	48.80 (0.99)	49.00 (1.49)	0.01	50.00 (0.00)	50.00 (0.00)	0.01	50.00 (0.00)	50.00 (0.00)	0.01
nBayes	0.00 (0.00)	0.00 (0.00)	0.04	0.00 (0.00)	0.00 (0.00)	0.14	0.00 (0.00)	0.00 (0.00)	0.70
AdaBoost	4.05 (1.72)	4.00 (1.49)	4.59	1.58 (1.38)	1.00 (1.68)	7.60	1.30 (1.22)	1.00 (1.49)	32.54
GBDT	3.43 (1.82)	3.00 (2.24)	4.18	0.65 (1.12)	0.00 (0.75)	4.86	0.05 (0.22)	0.00 (0.00)	18.58
XGBoost	30.00 (3.30)	30.00 (3.17)	0.01	29.68 (4.71)	29.00 (3.92)	0.02	31.00 (4.66)	30.50 (3.92)	0.02
$n = 100$ (balanced)									
Unimodal(ROT)	1.06 (0.35)	1.00 (0.37)	0.84	0.00 (0.00)	0.00 (0.00)	2.54	0.00 (0.00)	0.00 (0.00)	11.07
Unimodal(NS)	1.11 (0.36)	1.00 (0.19)	0.76	0.00 (0.00)	0.00 (0.00)	2.13	0.00 (0.00)	0.00 (0.00)	10.02
Quantile-mode	0.13 (0.17)	0.00 (0.19)	3.19	0.00 (0.00)	0.00 (0.00)	5.50	0.00 (0.00)	0.00 (0.00)	10.29
Naive-mode	12.77 (1.89)	13.00 (1.77)	0.06	4.34 (1.28)	4.25 (1.49)	0.08	0.73 (0.36)	0.75 (0.37)	0.18
Centroid	45.35 (2.35)	45.25 (1.96)	0.01	41.63 (1.82)	41.50 (1.40)	0.01	40.46 (1.71)	40.50 (1.91)	0.01
Median	11.36 (1.46)	11.50 (0.75)	0.04	3.58 (0.91)	3.75 (0.75)	0.04	0.27 (0.27)	0.25 (0.37)	0.24
Quantile	0.54 (0.55)	0.38 (0.37)	0.70	0.22 (0.47)	0.00 (0.19)	2.14	0.10 (0.15)	0.00 (0.19)	10.59
EQC	2.00 (0.00)	2.00 (0.00)	0.29	0.00 (0.00)	0.00 (0.00)	0.74	0.00 (0.00)	0.00 (0.00)	3.25
SVM	32.98 (2.37)	33.50 (2.71)	0.01	18.93 (1.49)	18.75 (1.54)	0.03	5.30 (1.01)	5.50 (1.12)	0.15
LDA	45.12 (2.33)	45.25 (2.01)	0.03	43.62 (1.99)	43.75 (1.49)	0.05	44.57 (2.06)	44.63 (1.73)	0.39
1-NN	46.22 (0.84)	46.13 (0.75)	0.01	50.00 (0.00)	50.00 (0.00)	0.01	50.00 (0.00)	50.00 (0.00)	0.02
nBayes	0.00 (0.00)	0.00 (0.00)	0.07	0.00 (0.00)	0.00 (0.00)	0.23	0.00 (0.00)	0.00 (0.00)	1.13
AdaBoost	0.35 (0.30)	0.25 (0.37)	4.66	0.04 (0.10)	0.00 (0.00)	8.67	0.01 (0.04)	0.00 (0.00)	36.30
GBDT	0.14 (0.18)	0.00 (0.19)	4.57	0.00 (0.00)	0.00 (0.00)	4.93	0.00 (0.00)	0.00 (0.00)	21.91
XGBoost	25.44 (2.74)	25.38 (3.03)	0.02	24.49 (2.35)	24.38 (1.77)	0.02	25.98 (2.21)	25.75 (2.33)	0.03
$n = 200$ (balanced)									
Unimodal(ROT)	1.06 (0.19)	1.10 (0.22)	0.84	0.00 (0.00)	0.00 (0.00)	2.45	0.00 (0.00)	0.00 (0.00)	11.23
Unimodal(NS)	1.11 (0.16)	1.10 (0.15)	0.83	0.00 (0.00)	0.00 (0.00)	2.42	0.00 (0.00)	0.00 (0.00)	10.07
Quantile-mode	0.32 (0.12)	0.30 (0.15)	0.15	0.00 (0.00)	0.00 (0.00)	0.41	0.00 (0.00)	0.00 (0.00)	1.84
Naive-mode	11.06 (1.23)	11.00 (1.40)	3.50	3.78 (0.51)	3.70 (0.45)	5.92	0.48 (0.22)	0.50 (0.24)	11.09
Centroid	44.51 (1.37)	44.55 (1.59)	0.01	43.28 (1.29)	43.50 (1.34)	0.01	39.05 (0.92)	39.00 (0.86)	0.02
Median	11.78 (0.75)	11.70 (0.69)	0.02	2.63 (0.52)	2.50 (0.47)	0.05	0.41 (0.17)	0.40 (0.15)	0.27
Quantile	0.43 (0.16)	0.40 (0.15)	0.81	0.07 (0.07)	0.10 (0.07)	2.74	0.06 (0.10)	0.00 (0.07)	11.80
EQC	1.30 (0.00)	1.30 (0.00)	0.30	0.00 (0.00)	0.00 (0.00)	0.80	0.00 (0.00)	0.00 (0.00)	3.63
SVM	33.39 (1.06)	33.20 (1.23)	0.02	25.85 (0.98)	25.80 (0.86)	0.05	14.35 (0.88)	14.35 (0.62)	0.27
LDA	44.36 (1.49)	44.50 (1.31)	0.02	44.44 (1.12)	44.60 (1.01)	0.08	42.15 (1.31)	41.90 (1.18)	0.80
1-NN	44.03 (0.74)	43.95 (0.67)	0.01	50.00 (0.00)	50.00 (0.00)	0.01	50.00 (0.00)	50.00 (0.00)	0.06
nBayes	0.01 (0.27)	0.00 (0.00)	0.12	0.00 (0.00)	0.00 (0.00)	0.44	0.00 (0.00)	0.00 (0.00)	2.27
AdaBoost	0.04 (0.08)	0.00 (0.07)	4.99	0.00 (0.02)	0.00 (0.00)	10.74	0.00 (0.00)	0.00 (0.00)	46.66
GBDT	0.01 (0.02)	0.00 (0.00)	5.84	0.00 (0.00)	0.00 (0.00)	8.55	0.00 (0.00)	0.00 (0.00)	38.55
XGBoost	23.57 (1.63)	24.00 (1.64)	0.02	24.14 (1.61)	24.15 (1.21)	0.02	24.41 (1.64)	24.30 (1.53)	0.03

Table S18: Misclassification rates multiplied by 100 for Case 7 of Example 1 with imbalanced setting

	$p = 50$			$p = 200$			$p = 1000$		
	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time
$n = 50$ (imbalanced)									
Unimodal(ROT)	2.08 (0.77)	2.00 (1.00)	0.78	0.00 (0.00)	0.00 (0.00)	2.23	0.00 (0.00)	0.00 (0.00)	10.81
Unimodal(NS)	2.53 (0.90)	2.67 (0.50)	0.76	0.00 (0.00)	0.00 (0.00)	1.99	0.00 (0.00)	0.00 (0.00)	9.91
Quantile-mode	0.53 (0.43)	0.67 (0.50)	3.07	0.00 (0.00)	0.00 (0.00)	5.26	0.00 (0.00)	0.00 (0.00)	9.83
Naïve-mode	12.57 (3.01)	12.00 (2.49)	0.07	7.87 (2.23)	7.33 (2.49)	0.10	4.88 (1.82)	5.33 (1.62)	0.15
Centroid	27.67 (2.41)	28.00 (2.61)	0.01	28.22 (1.44)	28.00 (1.49)	0.01	29.28 (1.64)	29.33 (1.99)	0.01
Median	12.20 (2.08)	12.67 (1.49)	0.02	5.65 (1.63)	6.00 (1.49)	0.07	2.20 (0.84)	2.00 (0.62)	0.22
Quantile	1.87 (1.45)	1.67 (1.12)	0.69	0.28 (0.45)	0.00 (0.50)	2.25	0.00 (0.00)	0.00 (0.00)	9.74
EQC	2.08 (0.53)	2.00 (0.00)	0.24	0.00 (0.00)	0.00 (0.00)	0.75	0.00 (0.00)	0.00 (0.00)	3.41
SVM	33.10 (0.41)	33.33 (0.50)	0.01	15.58 (2.93)	15.67 (2.24)	0.03	0.22 (0.35)	0.00 (0.50)	0.13
LDA	29.92 (2.95)	30.00 (2.99)	0.01	34.77 (3.32)	34.67 (2.99)	0.04	26.72 (2.27)	26.00 (2.11)	0.32
1-NN	31.62 (0.81)	32.00 (0.50)	0.01	33.33 (0.00)	33.33 (0.00)	0.01	33.33 (0.00)	33.33 (0.00)	0.01
nBayes	0.00 (0.00)	0.00 (0.00)	0.06	0.00 (0.00)	0.00 (0.00)	0.18	0.00 (0.00)	0.00 (0.00)	0.90
AdaBoost	2.63 (1.42)	2.67 (1.49)	4.21	1.53 (1.23)	1.33 (1.00)	8.12	1.43 (1.02)	1.33 (1.00)	33.41
GBDT	1.27 (1.02)	1.33 (1.00)	4.59	0.08 (0.22)	0.00 (0.00)	4.72	0.00 (0.00)	0.00 (0.00)	18.74
XGBoost	23.10 (3.08)	23.00 (3.11)	0.01	24.85 (2.78)	24.00 (3.48)	0.02	24.48 (3.22)	24.33 (3.48)	0.02
$n = 100$ (imbalanced)									
Unimodal(ROT)	0.95 (0.28)	1.00 (0.25)	0.80	0.00 (0.00)	0.00 (0.00)	2.16	0.00 (0.00)	0.00 (0.00)	11.08
Unimodal(NS)	0.99 (0.27)	1.00 (0.25)	0.68	0.00 (0.00)	0.00 (0.00)	1.87	0.00 (0.00)	0.00 (0.00)	9.82
Quantile-mode	0.20 (0.16)	0.17 (0.25)	3.17	0.00 (0.00)	0.00 (0.00)	5.62	0.00 (0.00)	0.00 (0.00)	10.67
Naive-mode	7.75 (1.02)	7.83 (0.90)	0.06	3.11 (0.72)	3.08 (0.87)	0.08	0.78 (0.32)	0.67 (0.37)	0.18
Centroid	37.85 (1.86)	38.58 (1.93)	0.01	32.58 (1.27)	32.33 (1.06)	0.01	28.15 (0.75)	28.17 (0.87)	0.01
Median	8.60 (1.10)	8.42 (1.37)	0.02	2.35 (0.63)	2.33 (0.65)	0.06	0.44 (0.17)	0.50 (0.16)	0.25
Quantile	0.61 (0.36)	0.50 (0.04)	0.77	0.10 (0.23)	0.00 (0.12)	2.29	0.07 (0.13)	0.00 (0.12)	10.47
EQC	2.15 (0.11)	2.17 (0.00)	0.25	0.16 (0.03)	0.17 (0.00)	0.81	0.00 (0.00)	0.00 (0.00)	3.61
SVM	33.19 (0.23)	33.33 (0.12)	0.01	20.66 (1.63)	20.58 (1.90)	0.04	6.91 (0.90)	7.00 (0.87)	.021
LDA	27.90 (1.12)	27.75 (1.37)	0.01	32.27 (1.60)	3.22 (1.96)	0.06	36.64 (2.52)	36.67 (2.24)	0.49
1-NN	31.44 (0.44)	31.50 (0.37)	0.01	33.33 (0.00)	33.33 (0.00)	0.01	33.33 (0.00)	33.33 (0.00)	0.03
nBayes	0.10 (0.11)	0.08 (0.12)	0.10	0.00 (0.00)	0.00 (0.00)	0.33	0.00 (0.00)	0.00 (0.00)	1.71
AdaBoost	0.40 (0.22)	0.33 (0.16)	5.03	0.06 (0.11)	0.00 (0.12)	9.50	0.02 (0.05)	0.00 (0.00)	40.62
GBDT	0.12 (0.11)	0.17 (0.12)	4.80	0.00 (0.00)	0.00 (0.00)	6.79	0.00 (0.00)	0.00 (0.00)	28.31
XGBoost	21.73 (1.48)	22.08 (1.80)	0.01	21.79 (1.59)	21.67 (1.90)	0.02	22.42 (1.44)	22.50 (1.68)	0.03
$n = 200$ (imbalanced)									
Unimodal(ROT)	0.94 (0.16)	0.93 (0.16)	0.79	0.00 (0.00)	0.00 (0.00)	2.27	0.00 (0.00)	0.00 (0.00)	11.90
Unimodal(NS)	0.90 (0.18)	0.90 (0.15)	0.73	0.00 (0.00)	0.00 (0.00)	2.10	0.00 (0.00)	0.00 (0.00)	10.68
Quantile-mode	0.36 (0.11)	0.40 (0.15)	3.79	0.00 (0.00)	0.00 (0.00)	6.45	0.00 (0.00)	0.00 (0.00)	12.22
Naive-mode	8.16 (0.79)	8.13 (0.78)	0.07	2.04 (0.37)	2.07 (0.36)	0.13	0.38 (0.15)	0.40 (0.15)	0.23
Centroid	43.01 (1.16)	43.03 (1.29)	0.01	35.48 (0.95)	35.50 (1.02)	0.01	30.86 (0.78)	30.83 (0.77)	0.02
Median	9.40 (0.45)	9.37 (0.45)	0.02	1.81 (0.27)	1.87 (0.25)	0.08	0.21 (0.09)	0.20 (0.10)	0.23
Quantile	0.40 (0.17)	0.40 (0.15)	0.92	0.05 (0.07)	0.00 (0.05)	3.01	0.00 (0.00)	0.00 (0.00)	11.49
EQC	1.01 (0.04)	1.00 (0.00)	0.32	0.00 (0.00)	0.00 (0.00)	0.96	0.00 (0.00)	0.00 (0.00)	4.49
SVM	33.33 (0.01)	33.33 (0.00)	0.03	22.58 (1.04)	22.47 (1.13)	0.08	13.90 (0.87)	14.03 (0.70)	0.47
LDA	28.72 (0.54)	28.70 (0.51)	0.02	28.31 (0.99)	28.17 (1.02)	0.11	31.46 (0.91)	31.63 (0.90)	1.04
1-NN	30.50 (0.42)	30.43 (0.40)	0.01	33.33 (0.00)	33.33 (0.00)	0.02	33.33 (0.00)	33.33 (0.00)	0.12
nBayes	0.02 (0.03)	0.00 (0.05)	0.16	0.00 (0.00)	0.00 (0.00)	0.62	0.00 (0.00)	0.00 (0.00)	3.07
AdaBoost	0.04 (0.04)	0.00 (0.05)	5.58	0.00 (0.01)	0.00 (0.00)	13.26	0.00 (0.00)	0.00 (0.00)	57.15
GBDT	0.00 (0.00)	0.00 (0.00)	5.87	0.00 (0.00)	0.00 (0.00)	20.76	0.00 (0.00)	0.00 (0.00)	51.98
XGBoost	20.63 (1.28)	20.67 (1.28)	0.02	21.07 (1.08)	21.07 (1.06)	0.02	22.18 (1.11)	22.13 (0.86)	0.04

Table S19: Misclassification rates multiplied by 100 for Case 1 of Example 2 with balanced and imbalanced settings

	$n = 50$ (balanced)									$n = 50$ (imbalanced)								
	$p = 50$			$p = 200$			$p = 1000$			$p = 50$			$p = 200$			$p = 1000$		
	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time
Multimodal	0.50 (0.72)	0.00 (0.75)	0.94	0.00 (0.00)	0.00 (0.00)	2.99	0.00 (0.00)	0.00 (0.00)	13.58	0.00 (0.00)	0.00 (0.00)	0.91	0.00 (0.00)	0.00 (0.00)	3.06	0.00 (0.00)	0.00 (0.00)	13.71
Centroid	3.10 (1.28)	3.00 (1.49)	0.01	0.00 (0.00)	0.00 (0.00)	0.01	0.00 (0.00)	0.00 (0.00)	0.01	2.15 (1.10)	2.00 (1.00)	0.01	0.00 (0.00)	0.00 (0.00)	0.01	0.00 (0.00)	0.00 (0.00)	0.01
Median	16.35 (2.73)	16.00 (2.99)	0.01	2.30 (1.32)	2.00 (1.49)	0.04	0.15 (0.36)	0.00 (0.00)	0.20	16.82 (2.11)	16.67 (1.99)	0.01	3.37 (1.35)	3.33 (1.00)	0.05	0.27 (0.36)	0.00 (0.50)	0.18
Quantile	0.00 (0.00)	0.00 (0.00)	0.62	0.00 (0.00)	0.00 (0.00)	1.99	0.00 (0.00)	0.00 (0.00)	9.80	0.00 (0.00)	0.00 (0.00)	0.62	0.00 (0.00)	0.00 (0.00)	2.12	0.00 (0.00)	0.00 (0.00)	9.86
EQC	10.98 (0.16)	11.00 (0.00)	0.23	0.00 (0.00)	0.00 (0.00)	0.65	0.00 (0.00)	0.00 (0.00)	2.96	7.43 (0.63)	7.33 (0.00)	0.21	0.65 (0.11)	0.67 (0.00)	0.74	0.00 (0.00)	0.00 (0.00)	3.08
SVM	2.13 (1.32)	2.00 (1.49)	0.01	0.00 (0.00)	0.00 (0.00)	0.02	0.00 (0.00)	0.00 (0.00)	0.12	33.33 (0.00)	33.33 (0.00)	0.01	33.33 (0.00)	33.33 (0.00)	0.02	11.77 (3.02)	11.33 (2.49)	0.14
LDA	7.60 (2.48)	7.00 (2.24)	0.01	3.65 (2.30)	3.00 (2.24)	0.04	0.43 (0.64)	0.00 (0.75)	0.28	3.70 (1.47)	3.33 (1.49)	0.01	4.68 (2.40)	4.67 (2.99)	0.04	0.20 (0.38)	0.00 (0.12)	0.31
1-NN	13.08 (3.55)	13.00 (3.17)	0.01	0.95 (0.96)	1.00 (0.75)	0.01	0.03 (0.16)	0.00 (0.00)	0.01	11.52 (1.87)	11.33 (1.49)	0.01	0.48 (0.54)	0.67 (0.50)	0.01	0.00 (0.00)	0.00 (0.00)	0.01
nBayes	2.95 (1.52)	3.00 (1.49)	0.03	0.00 (0.00)	0.00 (0.00)	0.13	0.00 (0.00)	0.00 (0.00)	0.62	1.77 (1.03)	1.33 (1.12)	0.04	0.00 (0.00)	0.00 (0.00)	0.21	0.00 (0.00)	0.00 (0.00)	0.87
AdaBoost	0.15 (0.53)	0.00 (0.00)	4.23	0.00 (0.00)	0.00 (0.00)	7.48	0.00 (0.00)	0.00 (0.00)	32.71	0.18 (0.45)	0.00 (0.00)	4.12	0.07 (0.20)	0.00 (0.00)	7.88	0.00 (0.00)	0.00 (0.00)	33.62
GBDT	1.43 (1.34)	1.00 (1.49)	3.50	0.03 (0.16)	0.00 (0.00)	4.04	0.00 (0.00)	0.00 (0.00)	17.61	0.57 (0.68)	0.67 (0.50)	4.15	0.00 (0.00)	0.00 (0.00)	4.49	0.00 (0.00)	0.00 (0.00)	18.61
XGBoost	15.15 (5.57)	14.00 (3.17)	0.01	16.40 (4.09)	16.00 (3.92)	0.02	17.15 (5.54)	16.50 (5.97)	0.03	9.53 (2.62)	9.00 (1.99)	0.01	12.73 (3.04)	12.67 (2.99)	0.02	12.25 (2.30)	12.00 (2.49)	0.02
	$n = 100$ (balanced)									$n = 100$ (imbalanced)								
Multimodal	0.18 (0.16)	0.25 (0.19)	1.22	0.00 (0.00)	0.00 (0.00)	3.32	0.00 (0.00)	0.00 (0.00)	15.12	0.11 (0.08)	0.17 (0.12)	0.94	0.00 (0.00)	0.00 (0.00)	3.20	0.00 (0.00)	0.00 (0.00)	14.52
Centroid	1.73 (0.54)	1.75 (0.56)	0.01	0.00 (0.00)	0.00 (0.00)	0.01	0.00 (0.00)	0.00 (0.00)	0.01	2.80 (0.54)	2.83 (0.50)	0.01	0.00 (0.00)	0.00 (0.00)	0.01	0.00 (0.00)	0.00 (0.00)	0.01
Median	13.89 (1.06)	13.88 (1.12)	0.02	2.70 (0.61)	2.75 (0.56)	0.05	0.04 (0.10)	0.00 (0.00)	0.20	17.49 (1.26)	17.50 (0.62)	0.02	2.80 (0.50)	2.83 (0.50)	0.04	0.28 (0.20)	0.33 (0.12)	0.21
Quantile	0.06 (0.04)	0.00 (0.00)	0.71	0.00 (0.00)	0.00 (0.00)	2.28	0.00 (0.00)	0.00 (0.00)	10.12	0.00 (0.03)	0.00 (0.00)	0.72	0.00 (0.00)	0.00 (0.00)	2.19	0.00 (0.00)	0.00 (0.00)	10.25
EQC	6.79 (0.24)	6.75 (0.00)	0.23	0.01 (0.04)	0.00 (0.00)	0.80	0.00 (0.00)	0.00 (0.00)	3.28	5.01 (0.08)	5.00 (0.00)	0.59	0.16 (0.03)	0.17 (0.00)	0.75	0.00 (0.00)	0.00 (0.00)	3.53
SVM	1.79 (0.53)	1.75 (0.42)	0.01	0.00 (0.00)	0.00 (0.00)	0.03	0.00 (0.00)	0.00 (0.00)	0.16	33.33 (0.00)	33.33 (0.00)	0.01	0.03 (0.08)	0.00 (0.00)	0.03	0.00 (0.00)	0.00 (0.00)	0.16
LDA	2.16 (0.67)	2.13 (0.93)	0.01	0.23 (0.28)	0.25 (0.19)	0.05	0.43 (0.35)	0.50 (0.23)	0.36	3.00 (0.68)	2.92 (0.75)	0.01	0.04 (0.07)	0.00 (0.03)	0.06	0.58 (0.38)	0.50 (0.40)	0.49
1-NN	10.75(1.23)	10.75 (1.35)	0.01	0.81 (0.38)	0.75 (0.37)	0.01	0.00 (0.00)	0.00 (0.00)	0.02	12.60 (1.26)	12.50 (1.21)	0.01	0.58 (0.34)	0.50 (0.37)	0.01	0.00 (0.00)	0.00 (0.00)	0.03
nBayes	1.88 (0.54)	1.88 (0.56)	0.07	0.00 (0.00)	0.00 (0.00)	0.22	0.00 (0.00)	0.00 (0.00)	1.23	2.60 (0.51)	2.50 (0.53)	0.08	0.00 (0.00)	0.00 (0.00)	0.33	0.00 (0.00)	0.00 (0.00)	1.66
AdaBoost	0.00 (0.00)	0.00 (0.00)	4.88	0.00 (0.00)	0.00 (0.00)	8.52	0.00 (0.00)	0.00 (0.00)	37.55	0.00 (0.00)	0.00 (0.00)	4.56	0.00 (0.00)	0.00 (0.00)	9.30	0.00 (0.00)	0.00 (0.00)	40.76
GBDT	0.00 (0.00)	0.00 (0.00)	3.67	0.00 (0.00)	0.00 (0.00)	4.85	0.00 (0.00)	0.00 (0.00)	21.52	0.00 (0.00)	0.00 (0.00)	4.91	0.00 (0.00)	0.00 (0.00)	7.05	0.00 (0.00)	0.00 (0.00)	28.82
XGBoost	12.33(1.54)	12.38 (1.54)	0.01	12.81 (1.74)	12.50 (1.59)	0.02	13.18 (1.67)	13.25 (1.49)	0.03	8.85 (0.86)	8.83 (0.78)	0.01	8.77 (1.27)	9.08 (1.24)	0.02	8.54 (0.97)	8.50 (0.81)	0.02
	$n = 200$ (balanced)									$n = 200$ (imbalanced)								
Multimodal	0.07 (0.05)	0.10 (0.07)	0.99	0.00 (0.00)	0.00 (0.00)	3.39	0.00 (0.00)	0.00 (0.00)	14.97	0.01 (0.05)	0.07 (0.05)	0.97	0.00 (0.00)	0.00 (0.00)	3.48	0.00 (0.00)	0.00 (0.00)	18.04
Centroid	2.35 (0.41)	2.30 (0.37)	0.01	0.00 (0.00)	0.00 (0.00)	0.01	0.00 (0.00)	0.00 (0.00)	0.01	2.40 (0.36)	2.40 (0.36)	0.01	0.00 (0.00)	0.00 (0.00)	0.01	0.00 (0.00)	0.00 (0.00)	0.01
Median	17.34 (0.95)	17.40 (1.06)	0.02	2.89 (0.42)	2.80 (0.52)	0.04	0.10 (0.07)	0.10 (0.09)	0.21	16.95 (0.58)	16.87 (0.58)	0.02	2.85 (0.34)	2.87 (0.31)	0.07	0.12 (0.08)	0.13 (0.05)	0.25
Quantile	0.01 (0.03)	0.00 (0.00)	0.77	0.00 (0.00)	0.00 (0.00)	2.37	0.00 (0.00)	0.00 (0.00)	11.04	0.01 (0.02)	0.00 (0.00)	1.17	0.00 (0.00)	0.00 (0.00)	2.45	0.00 (0.00)	0.00 (0.00)	11.26
EQC	5.41 (0.05)	5.40 (0.00)	0.29	0.20 (0.03)	0.20 (0.00)	0.77	0.00 (0.00)	0.00 (0.00)	3.50	6.00 (0.04)	6.00 (0.00)	0.30	0.20 (0.00)	0.20 (0.00)	0.91	0.00 (0.00)	0.00 (0.00)	4.30
SVM	2.22 (0.34)	2.20 (0.24)	0.02	0.00 (0.00)	0.00 (0.00)	0.05	0.00 (0.00)	0.00 (0.00)	0.19	28.42 (0.85)	28.47 (0.72)	0.02	0.02 (0.03)	0.00 (0.00)	0.07	0.00 (0.00)	0.00 (0.00)	0.27
LDA	2.81 (0.43)	2.70 (0.45)	0.02	0.00 (0.02)	0.00 (0.00)	0.08	0.00 (0.00)	0.00 (0.00)	0.63	2.43 (0.26)	2.47 (0.26)	0.02	0.00 (0.01)	0.00 (0.00)	0.16	0.00 (0.00)	0.00 (0.00)	0.97
1-NN	12.30 (0.82)	12.25 (0.62)	0.01	0.64 (0.18)	0.60 (0.15)	0.01	0.00 (0.00)	0.00 (0.00)	0.05	10.41 (0.59)	10.37 (0.56)	0.01	0.62 (0.21)	0.60 (0.26)	0.02	0.00 (0.00)	0.00 (0.00)	0.10
nBayes	2.41 (0.44)	2.40 (0.39)	0.10	0.00 (0.00)	0.00 (0.00)	0.58	0.00 (0.00)	0.00 (0.00)	2.11	2.09 (0.28)	2.07 (0.31)	0.16	0.00 (0.00)	0.00 (0.00)	0.76	0.00 (0.00)	0.00 (0.00)	3.08
AdaBoost	0.00 (0.00)	0.00 (0.00)	4.83	0.00 (0.00)	0.00 (0.00)	10.25	0.00 (0.00)	0.00 (0.00)	46.11	0.00 (0.00)	0.00 (0.00)	5.26	0.00 (0.00)	0.00 (0.00)	12.40	0.00 (0.00)	0.00 (0.00)	54.68
GBDT	0.00 (0.00)	0.00 (0.00)	5.23	0.00 (0.00)	0.00 (0.00)	8.06	0.00 (0.00)	0.00 (0.00)	35.95	0.00 (0.00)	0.00 (0.00)	5.89	0.00 (0.00)	0.00 (0.00)	20.38	0.00 (0.00)	0.00 (0.00)	50.10
XGBoost	12.00 (1.03)	12.05 (0.75)	0.01	12.55 (0.93)	12.50 (0.75)	0.02	12.40 (0.93)	12.40 (0.71)	0.03	8.67 (0.62)	8.67 (0.52)	0.01	8.35 (0.70)	8.40 (0.81)	0.02	8.31 (0.45)	8.33 (0.42)	0.04

Table S20: Misclassification rates multiplied by 100 for Case 2 of Example 2 with balanced and imbalanced settings

	$n = 50$ (balanced)									$n = 50$ (imbalanced)								
	$p = 50$			$p = 200$			$p = 1000$			$p = 50$			$p = 200$			$p = 1000$		
	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time
Mulitmodal	0.50 (0.72)	0.00 (0.75)	0.64	0.00 (0.00)	0.00 (0.00)	1.96	0.00 (0.00)	0.00 (0.00)	8.99	0.00 (0.00)	0.00 (0.00)	0.65	0.00 (0.00)	0.00 (0.00)	2.02	0.00 (0.00)	0.00 (0.00)	8.97
Centroid	3.10 (1.28)	3.00 (1.49)	0.01	0.00 (0.00)	0.00 (0.00)	0.01	0.00 (0.00)	0.00 (0.00)	0.01	2.15 (1.10)	2.00 (1.00)	0.01	0.00 (0.00)	0.00 (0.00)	0.01	0.00 (0.00)	0.00 (0.00)	0.01
Median	16.35 (2.73)	16.00 (2.99)	0.01	2.30 (1.32)	2.00 (1.49)	0.04	0.15 (0.36)	0.00 (0.00)	0.19	16.82 (2.11)	16.67 (1.99)	0.01	3.37 (1.35)	3.33 (1.00)	0.05	0.27 (0.36)	0.00 (0.50)	0.20
Quantile	0.00 (0.00)	0.00 (0.00)	0.60	0.00 (0.00)	0.00 (0.00)	2.15	0.00 (0.00)	0.00 (0.00)	9.67	0.00 (0.00)	0.00 (0.00)	0.63	0.00 (0.00)	0.00 (0.00)	2.17	0.00 (0.00)	0.00 (0.00)	9.74
EQC	4.00 (0.00)	4.00 (0.00)	0.22	0.00 (0.00)	0.00 (0.00)	0.63	0.00 (0.00)	0.00 (0.00)	2.87	9.85 (0.95)	10.00 (0.00)	0.21	1.31 (0.11)	1.33 (0.00)	0.69	0.00 (0.00)	0.00 (0.00)	3.02
SVM	2.13 (1.32)	2.00 (1.49)	0.01	0.00 (0.00)	0.00 (0.00)	0.02	0.00 (0.00)	0.00 (0.00)	0.12	33.33 (0.00)	33.33 (0.00)	0.01	33.33 (0.00)	33.33 (0.00)	0.02	11.77 (3.02)	11.33 (2.49)	0.13
LDA	7.60 (2.48)	7.00 (2.24)	0.01	3.65 (2.30)	3.00 (2.24)	0.03	0.43 (0.64)	0.00 (0.75)	0.25	3.70 (1.47)	3.33 (1.49)	0.01	4.68 (2.40)	4.67 (2.99)	0.04	0.20 (0.38)	0.00 (0.12)	0.30
1-NN	13.08 (3.55)	13.00 (3.17)	0.01	0.95 (0.96)	1.00 (0.75)	0.01	0.03 (0.16)	0.00 (0.00)	0.01	11.52 (1.87)	11.33 (1.49)	0.01	0.48 (0.54)	0.67 (0.50)	0.01	0.00 (0.00)	0.00 (0.00)	0.01
nBayes	2.95 (1.52)	3.00 (1.49)	0.03	0.00 (0.00)	0.00 (0.00)	0.13	0.00 (0.00)	0.00 (0.00)	0.67	1.77 (1.03)	1.33 (1.12)	0.04	0.00 (0.00)	0.00 (0.00)	0.18	0.00 (0.00)	0.00 (0.00)	0.88
AdaBoost	0.15 (0.53)	0.00 (0.00)	4.02	0.00 (0.00)	0.00 (0.00)	7.40	0.00 (0.00)	0.00 (0.00)	31.95	0.18 (0.45)	0.00 (0.00)	4.11	0.07 (0.20)	0.00 (0.00)	8.17	0.00 (0.00)	0.00 (0.00)	33.54
GBDT	1.43 (1.34)	1.00 (1.49)	3.55	0.03 (0.16)	0.00 (0.00)	4.18	0.00 (0.00)	0.00 (0.00)	19.36	0.57 (0.68)	0.67 (0.50)	3.69	0.00 (0.00)	0.00 (0.00)	4.10	0.00 (0.00)	0.00 (0.00)	19.19
XGBoost	15.15 (5.57)	14.00 (3.17)	0.01	16.40 (4.09)	16.00 (3.92)	0.02	17.15 (5.54)	16.50 (5.97)	0.02	9.53 (2.62)	9.00 (1.99)	0.01	12.73 (3.04)	12.67 (2.99)	0.02	12.25 (2.30)	12.00 (2.49)	0.02
	$n = 100$ (balanced)									$n = 100$ (imbalanced)								
Multimodal	0.18 (0.16)	0.25 (0.19)	0.69	0.00 (0.00)	0.00 (0.00)	2.08	0.00 (0.00)	0.00 (0.00)	9.58	0.11 (0.08)	0.17 (0.12)	0.69	0.00 (0.00)	0.00 (0.00)	2.05	0.00 (0.00)	0.00 (0.00)	9.29
Centroid	1.73 (0.54)	1.75 (0.56)	0.01	0.00 (0.00)	0.00 (0.00)	0.01	0.00 (0.00)	0.00 (0.00)	0.01	2.80 (0.54)	2.83 (0.50)	0.01	0.00 (0.00)	0.00 (0.00)	0.01	0.00 (0.00)	0.00 (0.00)	0.01
Median	13.89 (1.06)	13.88 (1.12)	0.04	2.70 (0.61)	2.75 (0.56)	0.05	0.04 (0.10)	0.00 (0.00)	0.22	17.49 (1.26)	17.50 (0.62)	0.02	2.80 (0.50)	2.83 (0.50)	0.06	0.28 (0.20)	0.33 (0.12)	0.21
Quantile	0.06 (0.04)	0.00 (0.00)	0.63	0.00 (0.00)	0.00 (0.00)	2.14	0.00 (0.00)	0.00 (0.00)	9.84	0.00 (0.03)	0.00 (0.00)	0.73	0.00 (0.00)	0.00 (0.00)	2.21	0.00 (0.00)	0.00 (0.00)	10.46
EQC	3.21 (0.28)	3.25 (0.00)	0.23	0.25 (0.00)	0.25 (0.00)	0.69	0.00 (0.00)	0.00 (0.00)	3.08	4.78 (0.32)	4.83 (0.00)	0.25	0.00 (0.00)	0.00 (0.00)	0.71	0.00 (0.00)	0.00 (0.00)	3.42
SVM	1.79 (0.53)	1.75 (0.42)	0.01	0.00 (0.00)	0.00 (0.00)	0.04	0.00 (0.00)	0.00 (0.00)	0.13	33.33 (0.00)	33.33 (0.00)	0.01	0.03 (0.08)	0.00 (0.00)	0.03	0.00 (0.00)	0.00 (0.00)	0.17
LDA	2.16 (0.67)	2.13 (0.93)	0.01	0.23 (0.28)	0.25 (0.19)	0.05	0.43 (0.35)	0.50 (0.23)	0.38	3.00 (0.68)	2.92 (0.75)	0.02	0.04 (0.07)	0.00 (0.03)	0.05	0.58 (0.38)	0.50 (0.40)	0.49
1-NN	10.75(1.23)	10.75 (1.35)	0.01	0.81 (0.38)	0.75 (0.37)	0.01	0.00 (0.00)	0.00 (0.00)	0.02	12.60 (1.26)	12.50 (1.21)	0.01	0.58 (0.34)	0.50 (0.37)	0.01	0.00 (0.00)	0.00 (0.00)	0.03
nBayes	1.88 (0.54)	1.88 (0.56)	0.06	0.00 (0.00)	0.00 (0.00)	0.23	0.00 (0.00)	0.00 (0.00)	1.18	2.60 (0.51)	2.50 (0.53)	0.08	0.00 (0.00)	0.00 (0.00)	0.32	0.00 (0.00)	0.00 (0.00)	1.61
AdaBoost	0.00 (0.00)	0.00 (0.00)	4.15	0.00 (0.00)	0.00 (0.00)	8.93	0.00 (0.00)	0.00 (0.00)	35.99	0.00 (0.00)	0.00 (0.00)	4.42	0.00 (0.00)	0.00 (0.00)	9.79	0.00 (0.00)	0.00 (0.00)	41.36
GBDT	0.00 (0.00)	0.00 (0.00)	3.89	0.00 (0.00)	0.00 (0.00)	4.75	0.00 (0.00)	0.00 (0.00)	21.67	0.00 (0.00)	0.00 (0.00)	4.49	0.00 (0.00)	0.00 (0.00)	6.47	0.00 (0.00)	0.00 (0.00)	28.52
XGBoost	12.33(1.54)	12.38 (1.54)	0.01	12.81 (1.74)	12.50 (1.59)	0.02	13.18 (1.67)	13.25 (1.49)	0.02	8.85 (0.86)	8.83 (0.78)	0.01	8.77 (1.27)	9.08 (1.24)	0.02	8.54 (0.97)	8.50 (0.81)	0.03
	$n = 200$ (balanced)									$n = 200$ (imbalanced)								
Multimodal	0.07 (0.05)	0.10 (0.07)	0.70	0.00 (0.00)	0.00 (0.00)	2.34	0.00 (0.00)	0.00 (0.00)	10.29	0.01 (0.05)	0.07 (0.05)	0.77	0.00 (0.00)	0.00 (0.00)	2.36	0.00 (0.00)	0.00 (0.00)	10.64
Centroid	2.35 (0.41)	2.30 (0.37)	0.01	0.00 (0.00)	0.00 (0.00)	0.01	0.00 (0.00)	0.00 (0.00)	0.01	2.40 (0.36)	2.40 (0.36)	0.01	0.00 (0.00)	0.00 (0.00)	0.01	0.00 (0.00)	0.00 (0.00)	0.02
Median	17.34 (0.95)	17.40 (1.06)	0.02	2.89 (0.42)	2.80 (0.52)	0.04	0.10 (0.07)	0.10 (0.09)	0.23	16.95 (0.58)	16.87 (0.58)	0.02	2.85 (0.34)	2.87 (0.31)	0.07	0.12 (0.08)	0.13 (0.05)	0.21
Quantile	0.01 (0.03)	0.00 (0.00)	0.75	0.00 (0.00)	0.00 (0.00)	2.35	0.00 (0.00)	0.00 (0.00)	10.51	0.01 (0.02)	0.00 (0.00)	0.86	0.00 (0.00)	0.00 (0.00)	2.51	0.00 (0.00)	0.00 (0.00)	11.27
EQC	2.1 (0.04)	2.1 (0.00)	0.28	0.01 (0.04)	0.00 (0.00)	0.74	0.00 (0.00)	0.00 (0.00)	3.52	4.12 (0.06)	4.13 (0.00)	0.28	0.33 (0.05)	0.33 (0.00)	0.89	0.00 (0.00)	0.00 (0.00)	4.34
SVM	2.22 (0.34)	2.20 (0.24)	0.02	0.00 (0.00)	0.00 (0.00)	0.04	0.00 (0.00)	0.00 (0.00)	0.25	28.42 (0.85)	28.47 (0.72)	0.02	0.02 (0.03)	0.00 (0.00)	0.07	0.00 (0.00)	0.00 (0.00)	0.43
LDA	2.81 (0.43)	2.70 (0.45)	0.02	0.00 (0.02)	0.00 (0.00)	0.08	0.00 (0.00)	0.00 (0.00)	0.60	2.43 (0.26)	2.47 (0.26)	0.02	0.00 (0.01)	0.00 (0.00)	0.09	0.00 (0.00)	0.00 (0.00)	0.94
1-NN	12.30 (0.82)	12.25 (0.62)	0.01	0.64 (0.18)	0.60 (0.15)	0.01	0.00 (0.00)	0.00 (0.00)	0.05	10.41 (0.59)	10.37 (0.56)	0.01	0.62 (0.21)	0.60 (0.26)	0.02	0.00 (0.00)	0.00 (0.00)	0.11
nBayes	2.41 (0.44)	2.40 (0.39)	0.11	0.00 (0.00)	0.00 (0.00)	0.43	0.00 (0.00)	0.00 (0.00)	2.12	2.09 (0.28)	2.07 (0.31)	0.16	0.00 (0.00)	0.00 (0.00)	0.60	0.00 (0.00)	0.00 (0.00)	3.01
AdaBoost	0.00 (0.00)	0.00 (0.00)	4.87	0.00 (0.00)	0.00 (0.00)	11.09	0.00 (0.00)	0.00 (0.00)	47.10	0.00 (0.00)	0.00 (0.00)	5.72	0.00 (0.00)	0.00 (0.00)	13.65	0.00 (0.00)	0.00 (0.00)	60.02
GBDT	0.00 (0.00)	0.00 (0.00)	4.64	0.00 (0.00)	0.00 (0.00)	7.67	0.00 (0.00)	0.00 (0.00)	36.74	0.00 (0.00)	0.00 (0.00)	5.70	0.00 (0.00)	0.00 (0.00)	19.95	0.00 (0.00)	0.00 (0.00)	51.34
XGBoost	12.00 (1.03)	12.05 (0.75)	0.01	12.55 (0.93)	12.50 (0.75)	0.02	12.40 (0.93)	12.40 (0.71)	0.03	8.67 (0.62)	8.67 (0.52)	0.01	8.35 (0.70)	8.40 (0.81)	0.02	8.31 (0.45)	8.33 (0.42)	0.04

Table S21: Misclassification rates multiplied by 100 for Case 3 of Example 2 with balanced and imbalanced settings

	$n = 50$ (balanced)									$n = 50$ (imbalanced)								
	$p = 50$			$p = 200$			$p = 1000$			$p = 50$			$p = 200$			$p = 1000$		
	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time
Multimodal	0.16 (5.84)	15.00 (4.66)	0.68	10.98 (3.63)	11.00 (3.73)	2.02	1.70 (1.11)	2.00 (0.93)	9.01	21.30 (2.38)	22.00 (2.61)	0.69	7.25 (2.39)	7.33 (2.99)	2.08	3.65 (1.10)	4.00 (1.49)	9.13
Centroid	0.46 (3.53)	45.00 (3.17)	0.01	47.03 (4.00)	47.00 (1.87)	0.01	38.78 (3.87)	39.00 (4.66)	0.01	42.15 (2.86)	42.00 (2.61)	0.01	41.12 (2.73)	41.00 (3.48)	0.01	35.00 (3.05)	34.67 (2.49)	0.01
Median	0.39 (5.19)	37.50 (4.66)	0.01	37.40 (4.42)	38.00 (3.17)	0.05	25.68 (4.07)	26.00 (4.48)	0.19	35.95 (3.56)	36.00 (3.23)	0.01	29.88 (3.48)	30.33 (3.73)	0.05	19.18 (3.80)	19.33 (3.61)	0.18
Quantile	0.31 (4.99)	31.00 (4.66)	0.58	21.20 (5.67)	21.00 (3.92)	2.11	11.35 (2.83)	11.00 (2.99)	9.65	32.80 (3.83)	33.00 (3.48)	0.62	18.02 (4.07)	17.67 (3.61)	2.20	15.37 (3.10)	15.67 (2.99)	9.86
EQC	0.43 (1.11)	43.00 (0.00)	0.24	25.15 (0.95)	25.00 (0.00)	0.63	18.20 (1.26)	18.00 (0.00)	2.85	31.90 (0.63)	32.00 (0.00)	0.22	27.88 (0.74)	28.00 (0.00)	0.67	16.78 (0.63)	16.67 (0.00)	3.10
SVM	0.18 (3.65)	18.00 (3.17)	0.01	0.00 (0.00)	0.00 (0.00)	0.02	0.00 (0.00)	0.00 (0.00)	0.12	17.22 (2.12)	16.67 (1.99)	0.01	0.00 (0.00)	0.00 (0.00)	0.02	0.00 (0.00)	0.00 (0.00)	0.12
LDA	0.49 (4.48)	48.00 (3.73)	0.01	48.50 (3.88)	49.00 (4.48)	0.03	45.73 (4.76)	45.50 (4.48)	0.25	40.27 (3.18)	40.33 (2.61)	0.01	45.98 (4.35)	45.00 (4.60)	0.04	36.18 (2.89)	36.00 (2.24)	0.29
1NN	0.45 (4.08)	45.00 (4.66)	0.01	49.83 (4.47)	50.50 (3.73)	0.01	45.48 (4.94)	46.00 (5.41)	0.01	40.30 (3.79)	40.67 (3.98)	0.01	42.73 (3.41)	43.33 (3.73)	0.01	42.97 (3.59)	42.33 (3.98)	0.01
nBayes	0.46 (3.38)	46.00 (2.99)	0.03	47.53 (3.84)	48.00 (3.73)	0.13	39.55 (3.84)	39.00 (3.73)	0.63	37.68 (2.96)	38.00 (3.11)	0.05	37.22 (2.25)	37.33 (2.49)	0.18	33.37 (1.77)	33.33 (1.99)	0.80
GBDT	0.35 (3.27)	35.00 (2.99)	3.61	36.05 (3.96)	35.50 (3.92)	4.41	26.78 (5.43)	27.00 (5.97)	17.51	33.07 (5.98)	32.67 (6.97)	3.70	31.37 (4.89)	31.67 (5.10)	4.27	23.57 (4.39)	23.33 (4.10)	18.13
AdaBoost	0.35 (4.28)	35.00 (4.48)	4.11	34.55 (4.33)	35.00 (4.66)	7.47	28.28 (4.48)	28.00 (3.17)	31.25	30.23 (2.63)	30.00 (2.11)	4.11	25.82 (2.56)	25.33 (2.99)	8.09	24.62 (2.30)	24.67 (2.11)	33.49
XGBoost	0.44 (4.43)	44.00 (3.73)	0.01	47.35 (4.25)	47.00 (3.92)	0.02	42.95 (5.51)	43.00 (5.41)	0.02	36.42 (4.30)	36.00 (4.60)	0.01	38.18 (4.39)	38.00 (2.61)	0.02	37.58 (5.44)	37.33 (6.47)	0.02
	$n = 100$ (balanced)									$n = 100$ (imbalanced)								
Multimodal	10.95 (1.28)	11.00 (1.31)	0.75	9.02 (1.68)	8.75 (1.54)	2.13	0.00 (0.00)	0.00 (0.00)	9.32	18.80 (1.02)	19.00 (1.00)	0.76	8.07 (1.21)	8.00 (1.15)	2.22	1.25 (0.39)	1.17 (0.37)	9.84
Centroid	45.64 (1.70)	46.00 (1.87)	0.01	41.45 (2.06)	41.13 (2.10)	0.01	34.28 (2.55)	34.13 (2.89)	0.01	41.96 (1.70)	42.42 (1.52)	0.01	38.76 (1.58)	39.00 (1.87)	0.01	32.53 (1.53)	32.67 (1.27)	0.01
Median	33.29 (1.72)	33.50 (11.70)	0.02	19.82 (1.87)	19.88 (1.87)	0.04	9.55 (1.56)	9.50 (9.56)	0.18	31.30 (1.51)	31.42 (1.24)	0.02	18.95 (1.27)	18.75 (0.75)	0.04	8.50 (0.88)	8.50 (0.78)	0.19
Quantile	27.98 (2.00)	27.88 (1.87)	0.65	13.27 (1.59)	13.25 (1.39)	2.18	3.43 (1.28)	3.25 (0.98)	9.52	25.78 (1.75)	25.50 (1.40)	0.72	10.95 (1.18)	11.00 (1.27)	2.29	3.12 (0.64)	3.08 (0.40)	10.18
EQC	36.03 (0.20)	36.00 (0.00)	0.25	25.68 (0.47)	25.75 (0.00)	0.75	10.95 (0.12)	11.00 (0.00)	3.02	29.58 (0.53)	29.76 (0.00)	0.24	17.00 (0.00)	17.00 (0.00)	0.73	10.13 (0.24)	10.17 (0.00)	3.47
SVM	32.04 (1.78)	32.00 (1.59)	0.01	9.90 (1.31)	9.88 (1.17)	0.03	0.00 (0.00)	0.00 (0.00)	0.15	30.57 (2.29)	30.08 (3.23)	0.01	13.85 (1.23)	14.00 (1.40)	0.03	0.00 (0.00)	0.00 (0.00)	0.22
LDA	45.36 (1.91)	45.25 (2.10)	0.01	42.40 (2.53)	42.50 (2.66)	0.05	43.79 (2.84)	43.75 (2.75)	0.35	35.79 (1.60)	35.75 (1.43)	0.01	38.09 (1.50)	38.00 (1.03)	0.06	42.75 (1.97)	42.75 (1.49)	0.49
1NN	48.37 (2.13)	48.13 (2.15)	0.01	49.56 (2.34)	49.75 (2.38)	0.01	46.11 (2.16)	46.25 (1.87)	0.01	42.58 (1.53)	42.67 (1.59)	0.01	41.71 (2.02)	41.67 (2.18)	0.01	42.63 (2.21)	42.42 (2.24)	0.03
nBayes	45.66 (2.01)	45.50 (1.87)	0.06	41.13 (1.95)	41.25 (1.35)	0.25	34.97 (2.24)	34.88 (2.24)	1.10	35.57 (1.47)	35.50 (1.49)	0.09	34.30 (1.55)	34.33 (1.55)	0.33	31.68 (1.41)	31.58 (1.59)	1.64
GBDT	24.43 (1.75)	24.75 (1.35)	4.19	15.39 (1.57)	15.25 (1.87)	4.62	9.13 (1.41)	9.13 (1.17)	21.63	21.67 (2.46)	21.75 (2.52)	4.50	15.20 (1.63)	14.83 (1.90)	6.20	10.41 (1.34)	10.08 (1.06)	27.62
AdaBoost	26.16 (1.51)	26.25 (0.98)	4.45	18.26 (1.78)	18.00 (1.82)	9.00	13.07 (1.61)	13.25 (1.31)	36.08	21.53 (1.21)	21.50 (1.03)	4.65	16.82 (1.02)	17.00 (1.00)	10.18	14.74 (1.51)	14.83 (1.65)	41.96
XGBoost	44.26 (2.51)	44.63 (2.15)	0.01	43.69 (2.16)	43.38 (1.96)	0.02	43.53 (2.41)	43.75 (2.24)	0.02	34.97 (1.57)	35.33 (1.49)	0.01	35.62 (1.92)	35.25 (2.18)	0.02	35.83 (1.85)	35.83 (1.40)	0.03
	$n = 200$ (balanced)									$n = 200$ (imbalanced)								
Multimodal	17.00 (0.77)	16.95 (0.67)	0.78	1.94 (0.41)	1.95 (0.32)	2.26	0.08 (0.07)	0.10 (0.07)	10.25	18.59 (0.72)	18.53 (0.76)	0.86	7.13 (0.52)	7.13 (0.52)	2.81	3.46 (0.38)	3.40 (0.36)	11.75
Centroid	44.35 (1.14)	44.40 (0.90)	0.01	37.72 (1.30)	37.80 (1.36)	0.01	30.33 (1.26)	30.45 (1.59)	0.01	42.11 (1.10)	42.03 (0.97)	0.01	35.54 (1.11)	35.60 (1.26)	0.01	28.70 (1.02)	28.73 (1.06)	0.02
Median	32.68 (1.26)	32.50 (1.44)	0.02	18.20 (1.18)	18.30 (0.99)	0.04	6.86 (0.70)	6.80 (0.63)	0.20	31.38 (1.00)	31.40 (1.11)	0.02	16.45 (0.79)	16.50 (0.56)	0.05	6.18 (0.51)	6.20 (0.51)	0.23
Quantile	28.17 (1.34)	28.20 (1.08)	0.77	10.45 (1.02)	10.40 (0.97)	2.59	2.25 (0.44)	2.20 (0.39)	10.23	25.46 (1.21)	25.27 (1.14)	0.86	8.74 (0.68)	8.73 (0.72)	2.56	1.95 (0.37)	1.93 (0.31)	11.02
EQC	29.52 (0.09)	29.50 (0.00)	0.22	19.64 (0.38)	19.70 (0.00)	0.78	7.21 (0.08)	7.20 (0.00)	3.46	26.02 (0.12)	26.00 (0.00)	0.24	17.23 (0.26)	17.27 (0.00)	0.90	6.62 (0.12)	6.60 (0.00)	4.38
SVM	37.77 (1.21)	37.65 (1.21)	0.02	22.68 (1.19)	22.70 (1.03)	0.04	3.15 (0.78)	3.05 (0.80)	0.22	33.33 (0.00)	33.33 (0.00)	0.02	22.40 (0.78)	22.47 (0.67)	0.07	7.54 (0.89)	7.57 (0.81)	0.41
LDA	44.70 (1.04)	44.60 (0.99)	0.02	39.09 (1.44)	38.95 (1.31)	0.07	36.34 (1.28)	36.10 (1.46)	0.62	34.39 (0.91)	34.33 (0.81)	0.02	32.57 (1.03)	32.63 (0.72)	0.09	30.05 (1.34)	29.93 (1.58)	1.03
1NN	50.26 (1.34)	50.25 (1.32)	0.01	47.33 (1.71)	47.45 (1.51)	0.01	47.93 (1.25)	48.00 (1.14)	0.05	44.13 (1.28)	43.87 (1.28)	0.01	42.51 (1.46)	42.20 (1.03)	0.02	42.16 (0.78)	42.03 (0.68)	0.11
nBayes	44.55 (1.18)	44.50 (0.93)	0.12	38.02 (1.37)	38.35 (1.59)	0.40	30.95 (1.04)	30.85 (1.01)	2.03	34.04 (0.69)	34.07 (0.60)	0.16	31.83 (1.02)	31.93 (1.06)	0.60	27.07 (0.94)	27.20 (0.88)	2.99
GBDT	20.69 (1.12)	20.70 (0.93)	5.01	10.37 (0.89)	10.35 (0.84)	7.84	4.86 (0.66)	4.85 (0.63)	35.58	20.01 (1.27)	19.83 (1.07)	5.91	10.80 (1.06)	10.67 (0.80)	19.72	6.61 (0.79)	6.63 (0.71)	49.75
AdaBoost	23.08 (1.00)	23.00 (0.99)	5.26	12.84 (1.12)	12.65 (1.14)	11.29	7.59 (0.99)	7.60 (0.91)	49.95	19.16 (0.99)	19.33 (1.06)	5.99	11.48 (0.78)	11.43 (0.82)	15.08	9.01 (0.66)	8.97 (0.67)	60.10
XGBoost	44.74 (1.08)	44.65 (1.06)	0.01	43.85 (1.59)	43.80 (1.51)	0.02	43.66 (1.63)	43.85 (1.79)	0.03	33.80 (1.07)	33.67 (1.24)	0.01	35.04 (1.18)	34.93 (1.08)	0.03	34.02 (0.96)	34.00 (0.92)	0.04

Table S22: Misclassification rates multiplied by 100 for Case 4 of Example 2 with balanced and imbalanced settings

	$n = 50$ (balanced)									$n = 50$ (imbalanced)								
	$p = 50$			$p = 200$			$p = 1000$			$p = 50$			$p = 200$			$p = 1000$		
	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time
Multimodal	2.40 (1.53)	2.00 (1.68)	0.81	0.05 (0.22)	0.00 (0.00)	2.18	0.00 (0.00)	0.00 (0.00)	9.52	4.00 (1.29)	4.00 (1.00)	0.78	0.02 (0.11)	0.00 (0.00)	2.22	0.00 (0.00)	0.00 (0.00)	9.72
Centroid	43.70 (3.91)	43.50 (3.17)	0.01	26.65 (3.70)	26.50 (2.99)	0.01	22.73 (3.10)	22.00 (3.17)	0.01	39.18 (3.19)	39.33 (2.86)	0.01	24.12 (3.12)	23.33 (2.99)	0.01	15.93 (2.28)	16.00 (2.24)	0.01
Median	46.90 (4.18)	47.00 (3.92)	0.01	31.83 (3.88)	31.00 (3.17)	0.04	30.58 (3.73)	31.00 (3.73)	0.19	42.19 (3.72)	42.67 (2.99)	0.02	34.73 (3.05)	34.00 (3.48)	0.06	28.03 (2.75)	28.67 (2.49)	0.18
Quantile	20.33 (4.46)	20.00 (4.10)	0.60	7.23 (3.29)	7.00 (3.92)	2.10	4.63 (2.31)	5.00 (2.24)	9.29	16.46 (3.38)	15.33 (3.48)	0.64	3.58 (2.75)	2.67 (2.99)	2.10	1.40 (0.99)	1.33 (1.00)	9.64
EQC	41.83 (1.11)	42.00 (0.00)	0.20	26.15 (0.95)	26.00 (0.00)	0.63	13.90 (0.63)	14.00 (0.00)	2.90	35.13 (1.26)	35.33 (0.00)	0.24	19.98 (0.11)	20.00 (0.00)	0.65	15.80 (1.26)	16.00 (0.00)	3.10
SVM	16.60 (2.90)	16.00 (3.17)	0.01	0.00 (0.00)	0.00 (0.00)	0.02	0.00 (0.00)	0.00 (0.00)	0.13	15.85 (2.09)	15.33 (2.24)	0.01	0.00 (0.00)	0.00 (0.00)	0.03	0.00 (0.00)	0.00 (0.00)	0.12
LDA	44.78 (3.57)	45.00 (3.17)	0.01	38.85 (5.41)	38.00 (3.92)	0.03	38.23 (4.27)	38.00 (4.10)	0.25	36.55 (3.37)	36.00 (3.23)	0.01	36.02 (4.08)	35.67 (4.10)	0.04	31.12 (3.33)	31.33 (3.48)	0.32
1NN	48.03 (4.80)	47.50 (5.41)	0.01	41.78 (3.74)	42.00 (3.36)	0.01	40.05 (4.33)	39.00 (3.92)	0.01	40.55 (3.73)	40.67 (3.48)	0.01	36.82 (3.60)	36.33 (3.23)	0.01	34.02 (2.80)	34.00 (2.99)	0.01
nBayes	45.48 (3.72)	45.50 (3.36)	0.03	28.98 (3.84)	29.00 (2.99)	0.12	26.75 (3.32)	26.00 (3.73)	0.62	37.42 (2.75)	37.33 (3.23)	0.04	25.38 (2.98)	26.00 (2.99)	0.17	21.03 (2.44)	21.00 (2.11)	0.86
GBDT	37.50 (4.56)	37.50 (3.17)	3.61	19.23 (2.91)	18.50 (2.99)	4.06	19.00 (4.65)	19.00 (5.97)	19.07	29.63 (5.23)	28.67 (5.35)	4.03	17.02 (3.81)	16.67 (3.98)	4.31	10.57 (2.50)	10.33 (1.99)	18.67
AdaBoost	30.48 (4.39)	30.00 (4.48)	3.95	17.93 (3.41)	18.00 (2.99)	7.36	17.35 (4.20)	17.00 (3.92)	31.85	20.83 (2.56)	20.67 (2.86)	4.12	15.58 (2.85)	15.33 (2.61)	8.03	15.67 (2.66)	15.67 (2.49)	32.91
XGBoost	43.63 (5.41)	43.00 (5.97)	0.01	40.03 (4.71)	40.50 (4.48)	0.02	42.05 (5.10)	42.50 (5.41)	0.02	33.47 (3.10)	33.33 (2.99)	0.01	33.40 (3.28)	32.67 (3.98)	0.02	33.18 (4.10)	33.33 (4.10)	0.02
	$n = 100$ (balanced)									$n = 100$ (imbalanced)								
Multimodal	3.41 (0.96)	3.50 (0.84)	0.83	0.01 (0.04)	0.00 (0.00)	2.30	0.00 (0.00)	0.00 (0.00)	9.87	4.03 (0.63)	4.00 (0.62)	0.82	0.00 (0.00)	0.00 (0.00)	2.39	0.00 (0.00)	0.00 (0.00)	10.44
Centroid	34.71 (1.77)	34.75 (1.91)	0.01	21.63 (1.67)	21.75 (1.59)	0.01	11.20 (1.23)	11.00 (1.17)	0.01	32.38 (1.61)	32.33 (1.74)	0.01	20.61 (1.36)	20.75 (1.43)	0.01	9.85 (0.85)	9.67 (0.90)	0.01
Median	43.31 (1.82)	43.75 (1.77)	0.02	39.16 (1.61)	39.00 (1.73)	0.04	28.46 (1.62)	28.25 (1.40)	0.20	43.06 (1.87)	43.17 (1.87)	0.02	34.50 (1.81)	34.16 (1.99)	0.04	27.92 (1.58)	27.83 (1.24)	0.20
Quantile	12.76 (1.41)	12.75 (0.98)	0.65	1.32 (0.62)	1.25 (0.75)	2.22	0.36 (0.30)	0.25 (0.37)	10.05	9.86 (1.23)	9.75 (1.40)	0.75	1.01 (0.37)	1.00 (0.25)	2.13	0.23 (0.24)	0.17 (0.28)	9.87
EQC	28.31 (0.36)	28.25 (0.00)	0.25	13.29 (0.24)	13.25 (0.00)	0.74	2.53 (0.20)	2.50 (0.00)	3.09	25.78 (0.34)	25.83 (0.00)	0.29	15.80 (0.24)	15.83 (0.00)	0.70	5.61 (0.37)	5.67 (0.00)	3.46
SVM	25.66 (1.75)	25.50 (1.91)	0.01	2.76 (0.86)	2.63 (0.79)	0.02	0.00 (0.00)	0.00 (0.00)	0.15	23.95 (1.43)	23.83 (1.18)	0.01	5.45 (1.09)	5.83 (1.06)	0.03	0.00 (0.00)	0.00 (0.00)	0.18
LDA	35.81 (1.88)	35.75 (1.59)	0.02	28.85 (2.95)	28.13 (3.26)	0.06	30.73 (2.96)	30.88 (3.17)	0.36	29.69 (1.50)	29.67 (1.40)	0.02	24.30 (1.73)	23.83 (1.46)	0.06	30.20 (2.31)	30.67 (2.15)	0.47
1NN	47.28 (2.18)	47.25 (2.29)	0.01	41.39 (1.94)	41.75 (2.05)	0.01	38.73 (2.07)	39.00 (1.91)	0.01	39.79 (1.58)	40.08 (1.68)	0.01	36.30 (1.69)	36.17 (1.71)	0.01	33.40 (1.41)	33.25 (1.74)	0.03
nBayes	35.81 (1.80)	35.75 (1.73)	0.09	24.58 (2.05)	24.38 (1.96)	0.25	12.78 (1.59)	12.88 (1.35)	1.07	28.58 (1.47)	28.33 (1.37)	0.09	20.94 (1.42)	21.00 (1.43)	0.32	11.34 (1.02)	11.17 (1.03)	1.62
GBDT	10.93 (1.53)	11.13 (1.49)	3.99	2.49 (0.70)	2.50 (0.75)	4.01	0.63 (0.45)	0.50 (0.37)	20.83	8.49 (1.30)	8.42 (1.52)	4.31	2.83 (0.83)	2.58 (0.78)	6.39	0.68 (0.32)	0.67 (0.40)	27.81
AdaBoost	11.49 (1.45)	11.63 (1.54)	4.30	3.48 (1.00)	3.75 (1.12)	9.21	1.79 (0.77)	1.75 (0.79)	36.32	7.99 (1.04)	7.92 (1.03)	4.50	3.06 (0.68)	3.17 (0.65)	10.23	1.63 (0.60)	1.50 (0.78)	41.73
XGBoost	39.69 (2.30)	39.88 (2.15)	0.01	38.34 (3.17)	38.38 (3.87)	0.02	39.35 (2.98)	39.00 (2.80)	0.02	29.95 (1.89)	29.83 (1.77)	0.01	30.35 (2.34)	30.50 (2.33)	0.02	29.57 (1.70)	29.58 (1.27)	0.03
	$n = 200$ (balanced)									$n = 200$ (imbalanced)								
Multimodal	5.41 (0.60)	5.40 (0.56)	0.87	0.10 (0.09)	0.10 (0.07)	2.46	0.00 (0.00)	0.00 (0.00)	10.90	6.67 (0.58)	6.60 (0.56)	0.96	0.04 (0.05)	0.00 (0.05)	2.86	0.00 (0.00)	0.00 (0.00)	12.39
Centroid	33.65 (1.21)	33.55 (0.80)	0.01	19.95 (1.27)	20.10 (1.23)	0.01	7.94 (0.63)	7.80 (0.60)	0.01	33.81 (0.98)	33.63 (0.98)	0.01	18.22 (0.87)	18.40 (0.93)	0.01	7.94 (0.63)	7.80 (0.60)	0.02
Median	44.06 (1.18)	44.05 (1.08)	0.02	38.87 (1.63)	38.60 (1.62)	0.04	28.35 (0.97)	28.50 (0.95)	0.20	43.55 (1.18)	43.70 (0.88)	0.02	35.63 (1.08)	35.70 (0.82)	0.05	28.35 (0.97)	28.50 (0.95)	0.22
Quantile	11.62 (0.92)	11.50 (0.90)	0.76	0.63 (0.24)	0.60 (0.32)	2.21	0.09 (0.13)	0.00 (0.07)	10.15	11.66 (0.64)	11.53 (0.65)	1.01	0.70 (0.21)	0.73 (0.16)	2.47	0.09 (0.13)	0.00 (0.07)	10.92
EQC	25.71 (0.05)	25.70 (0.00)	0.22	10.91 (0.03)	10.90 (0.00)	0.75	2.61 (0.08)	2.60 (0.00)	3.46	24.08 (0.36)	24.13 (0.00)	0.24	10.43 (0.21)	10.47 (0.00)	0.89	2.47 (0.00)	2.47 (0.00)	4.29
SVM	28.78 (1.14)	28.50 (0.82)	0.02	10.52 (0.99)	10.60 (0.73)	0.04	0.00 (0.00)	0.00 (0.00)	0.24	26.08 (0.73)	26.03 (0.80)	0.03	10.88 (0.90)	10.93 (0.86)	0.06	0.00 (0.00)	0.00 (0.00)	0.39
LDA	33.82 (0.93)	33.85 (0.78)	0.02	22.25 (1.11)	21.90 (1.27)	0.08	15.32 (1.04)	15.40 (0.50)	0.63	28.30 (0.80)	28.40 (0.67)	0.02	18.73 (0.88)	18.63 (0.90)	0.09	15.32 (1.04)	15.40 (0.47)	0.93
1NN	45.56 (1.77)	45.85 (1.68)	0.01	43.04 (1.35)	43.10 (1.49)	0.01	37.42 (1.51)	37.30 (1.51)	0.05	41.06 (1.07)	41.20 (1.13)	0.01	36.23 (1.33)	36.17 (1.38)	0.02	37.42 (1.51)	37.30 (1.51)	0.11
nBayes	33.86 (1.31)	33.95 (1.18)	0.11	20.92 (1.24)	21.00 (1.34)	0.39	8.75 (0.68)	8.80 (0.49)	2.07	28.51 (0.78)	28.37 (0.75)	0.16	17.48 (0.68)	17.53 (0.61)	0.58	8.75 (0.68)	8.80 (0.49)	2.90
GBDT	6.11 (0.66)	6.15 (0.82)	4.96	0.65 (0.27)	0.60 (0.30)	8.35	0.06 (0.09)	0.00 (0.07)	35.41	6.71 (0.87)	6.67 (0.95)	5.45	1.58 (0.31)	1.60 (0.36)	19.65	0.06 (0.09)	0.00 (0.07)	49.60
AdaBoost	7.04 (0.90)	7.10 (0.84)	5.10	1.27 (0.36)	1.25 (0.37)	11.30	0.25 (0.20)	0.20 (0.15)	49.29	6.12 (0.65)	6.03 (0.71)	6.08	1.13 (0.23)	1.10 (0.21)	14.85	0.25 (0.20)	0.20 (0.15)	60.08
XGBoost	39.01 (2.22)	38.80 (3.02)	0.01	38.89 (2.17)	38.65 (2.31)	0.02	38.60 (2.74)	38.10 (3.17)	0.03	29.43 (1.08)	29.37 (0.88)	0.01	29.27 (1.15)	29.33 (1.06)	0.02	38.60 (2.74)	38.10 (3.17)	0.04

Table S23: Misclassification rates multiplied by 100 for Case 5 of Example 2 with balanced and imbalanced settings

	$n = 50$ (balanced)									$n = 50$ (imbalanced)								
	$p = 50$			$p = 200$			$p = 1000$			$p = 50$			$p = 200$			$p = 1000$		
	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time
Multimodal	10.63 (6.08)	10.00 (3.73)	0.85	0.93 (1.83)	0.00 (0.75)	2.53	0.05 (0.22)	0.00 (0.00)	11.06	20.23 (5.13)	20.33 (4.73)	0.90	22.45 (4.38)	23.00 (5.10)	2.50	23.55 (5.30)	24.00 (5.22)	10.60
Centroid	35.20 (3.38)	35.00 (2.99)	0.01	27.75 (3.89)	27.50 (3.73)	0.01	13.78 (2.79)	14.00 (2.99)	0.01	33.37 (2.82)	33.33 (2.11)	0.01	30.05 (3.33)	29.67 (2.61)	0.01	14.58 (1.91)	14.67 (1.99)	0.01
Median	44.30 (3.64)	44.50 (3.17)	0.02	38.40 (3.71)	39.00 (3.36)	0.04	29.40 (3.25)	29.00 (3.17)	0.20	40.27 (3.69)	40.00 (2.99)	0.01	36.52 (3.30)	36.33 (3.11)	0.04	27.40 (3.14)	27.33 (2.99)	0.19
Quantile	20.28 (4.56)	20.50 (5.22)	0.61	5.18 (5.97)	4.00 (1.68)	2.00	0.33 (0.47)	0.00 (0.75)	9.56	18.68 (4.57)	18.67 (3.11)	0.66	5.60 (2.22)	5.00 (1.49)	2.12	1.20 (1.06)	1.33 (0.75)	9.60
EQC	35.98 (0.16)	36.00 (0.00)	0.20	32.78 (1.42)	33.00 (0.00)	0.64	13.88 (0.79)	14.00 (0.00)	3.14	33.97 (0.21)	34.00 (0.00)	0.26	27.38 (0.32)	27.33 (0.00)	0.76	12.82 (0.95)	12.67 (0.00)	3.14
SVM	12.55 (2.90)	12.00 (1.68)	0.01	0.00 (0.00)	0.00 (0.00)	0.02	0.00 (0.00)	0.00 (0.00)	0.11	15.18 (2.12)	15.33 (2.11)	0.01	0.02 (0.11)	0.00 (0.00)	0.02	0.00 (0.00)	0.00 (0.00)	0.15
LDA	39.15 (3.56)	39.00 (3.73)	0.01	37.20 (4.14)	37.00 (4.66)	0.03	37.75 (3.85)	38.00 (3.92)	0.26	33.87 (3.58)	33.33 (3.11)	0.01	41.57 (3.54)	43.00 (2.49)	0.04	29.38 (3.10)	29.33 (2.24)	0.32
1NN	46.53 (3.19)	47.00 (3.17)	0.01	41.93 (3.85)	42.00 (3.73)	0.01	33.83 (4.30)	34.00 (3.92)	0.01	42.77 (3.78)	42.67 (3.61)	0.01	39.63 (3.06)	39.33 (2.74)	0.01	33.90 (2.88)	34.33 (2.74)	0.01
nBayes	36.28 (3.23)	36.00 (3.73)	0.03	30.00 (3.80)	30.00 (3.73)	0.13	15.63 (2.82)	16.00 (2.24)	0.63	31.67 (2.84)	32.00 (2.61)	0.05	29.95 (2.68)	29.67 (2.99)	0.21	17.30 (2.25)	17.33 (1.99)	0.91
GBDT	33.20 (3.96)	33.00 (2.99)	3.50	27.30 (3.43)	28.00 (4.48)	4.25	20.93 (3.81)	21.00 (4.48)	18.98	30.30 (4.84)	30.00 (5.60)	4.00	30.65 (4.34)	30.00 (5.10)	4.20	19.90 (4.69)	19.33 (3.98)	19.09
AdaBoost	31.43 (3.64)	32.00 (3.92)	3.88	26.90 (5.36)	26.00 (5.97)	7.45	23.18 (4.33)	22.50 (4.48)	31.74	22.85 (2.70)	22.67 (2.99)	4.30	22.85 (2.42)	22.67 (2.11)	8.52	20.10 (2.10)	20.00 (1.99)	34.46
XGBoost	42.85 (4.37)	43.00 (4.85)	0.01	41.53 (5.12)	41.00 (5.41)	0.02	43.08 (4.73)	43.00 (4.66)	0.02	35.28 (5.01)	34.67 (4.73)	0.01	35.52 (4.57)	35.33 (5.60)	0.02	35.82 (4.17)	35.33 (3.61)	0.02
	$n = 100$ (balanced)									$n = 100$ (imbalanced)								
Multimodal	4.30 (1.31)	4.00 (1.35)	1.03	0.02 (0.07)	0.00 (0.00)	2.80	0.00 (0.00)	0.00 (0.00)	10.78	4.11 (0.84)	4.25 (0.87)	0.88	0.03 (0.06)	0.00 (0.00)	2.64	0.00 (0.00)	0.00 (0.00)	11.54
Centroid	35.09 (1.67)	35.25 (1.73)	0.01	20.93 (1.59)	21.00 (1.35)	0.01	10.56 (1.22)	10.75 (1.03)	0.01	34.55 (1.70)	34.83 (1.62)	0.01	19.93 (1.20)	19.83 (1.15)	0.01	7.08 (0.97)	6.92 (1.12)	0.01
Median	37.73 (2.18)	37.88 (1.77)	0.02	26.02 (1.85)	26.38 (1.59)	0.04	15.00 (1.42)	15.00 (1.49)	0.20	36.28 (2.16)	36.42 (1.90)	0.01	23.82 (1.29)	23.75 (1.15)	0.04	10.86 (1.25)	11.67 (1.27)	0.21
Quantile	16.92 (2.33)	16.88 (2.33)	0.66	2.47 (0.96)	2.25 (1.12)	2.18	0.24 (0.27)	0.25 (0.23)	10.41	15.62 (1.27)	15.33 (1.15)	0.70	1.62 (0.55)	1.50 (0.53)	2.23	0.05 (0.08)	0.00 (0.03)	10.36
EQC	29.94 (0.40)	30.00 (0.00)	0.51	16.27 (0.12)	16.25 (0.00)	0.75	6.30 (0.32)	6.25 (0.00)	3.13	24.98 (0.11)	25.00 (0.00)	0.22	16.80 (0.24)	16.83 (0.00)	0.74	5.74 (0.47)	5.67 (0.00)	3.66
SVM	25.64 (1.72)	25.50 (1.49)	0.01	2.78 (0.80)	2.75 (0.79)	0.03	0.00 (0.00)	0.00 (0.00)	0.16	24.43 (0.94)	24.50 (0.78)	0.01	5.66 (0.96)	5.50 (0.90)	0.03	0.00 (0.00)	0.00 (0.00)	0.19
LDA	36.01 (1.90)	36.00 (1.68)	0.02	28.20 (1.98)	28.25 (2.33)	0.05	30.73 (2.49)	30.50 (1.91)	0.57	30.98 (1.50)	30.83 (1.59)	0.01	23.98 (1.80)	24.17 (1.74)	0.06	26.85 (1.85)	26.92 (1.43)	0.48
1NN	47.88 (2.48)	47.38 (2.33)	0.01	42.82 (2.10)	42.75 (2.10)	0.01	34.26 (2.07)	34.25 (1.77)	0.03	43.00 (1.54)	42.92 (1.31)	0.01	37.53 (1.83)	37.42 (1.93)	0.01	31.72 (1.67)	31.83 (1.90)	0.03
nBayes	34.84 (1.69)	34.88 (15.86)	0.06	21.80 (1.84)	22.00 (1.77)	0.24	11.44 (1.29)	11.25 (1.03)	1.10	30.75 (1.36)	30.67 (1.30)	0.08	19.24 (1.13)	19.17 (1.18)	0.32	7.92 (0.97)	8.00 (1.00)	1.67
GBDT	14.50 (1.34)	14.50 (0.98)	3.82	5.51 (1.29)	5.50 (1.35)	4.81	2.28 (0.81)	2.25 (0.93)	23.38	12.46 (1.81)	12.50 (1.68)	4.56	4.52 (1.05)	4.42 (0.93)	6.37	1.60 (0.72)	1.33 (0.65)	29.10
AdaBoost	15.63 (1.48)	15.75 (1.96)	4.47	6.13 (1.22)	6.13 (1.31)	9.08	4.06 (1.15)	4.00 (1.12)	38.35	12.27 (1.39)	12.08 (1.03)	4.68	4.51 (0.83)	4.50 (0.75)	10.05	3.55 (0.80)	3.67 (0.65)	42.70
XGBoost	41.13 (2.14)	41.38 (2.01)	0.03	41.74 (2.41)	41.00 (2.24)	0.02	40.77 (3.01)	40.50 (2.15)	0.02	31.10 (1.56)	31.08 (1.15)	0.01	29.88 (2.22)	29.67 (1.77)	0.02	31.18 (1.76)	30.83 (1.31)	0.03
	$n = 200$ (balanced)									$n = 200$ (imbalanced)								
Multimodal	5.37 (0.63)	5.40 (0.60)	1.12	0.03 (0.05)	0.00 (0.07)	3.14	0.00 (0.00)	0.00 (0.00)	12.38	5.48 (0.59)	5.43 (0.46)	1.07	0.09 (0.06)	0.07 (0.05)	3.28	0.00 (0.00)	0.00 (0.00)	14.30
Centroid	33.57 (1.15)	33.60 (0.84)	0.01	18.44 (0.86)	18.30 (0.91)	0.01	7.48 (0.74)	7.40 (0.71)	0.01	32.31 (1.07)	32.43 (1.04)	0.01	17.28 (0.91)	17.27 (0.95)	0.01	7.65 (0.64)	7.70 (0.67)	0.02
Median	34.51 (1.24)	34.40 (1.23)	0.02	20.28 (0.98)	20.25 (0.73)	0.05	8.60 (0.83)	8.50 (0.67)	0.22	32.78 (0.90)	32.77 (1.06)	0.03	18.87 (0.91)	18.97 (0.81)	0.05	8.62 (0.61)	8.70 (0.46)	0.22
Quantile	14.62 (1.08)	14.50 (1.12)	0.85	1.75 (0.40)	1.70 (0.47)	2.28	0.07 (0.09)	0.00 (0.07)	10.94	14.76 (0.80)	14.73 (0.67)	0.90	1.61 (0.31)	1.60 (0.16)	2.49	0.08 (0.06)	0.07 (0.06)	11.27
EQC	27.35 (0.33)	27.30 (0.00)	0.29	13.53 (0.16)	13.50 (0.00)	0.87	4.91 (0.08)	4.80 (0.00)	3.49	24.75 (0.13)	24.73 (0.00)	0.23	13.59 (0.05)	13.60 (0.00)	0.88	4.87 (0.02)	4.87 (0.00)	4.10
SVM	29.25 (1.20)	29.25 (1.21)	0.02	9.29 (0.85)	9.30 (1.06)	0.04	0.00 (0.00)	0.00 (0.00)	0.22	25.86 (0.81)	25.93 (0.72)	0.02	9.91 (0.85)	9.90 (0.66)	0.07	0.01 (0.02)	0.00 (0.00)	0.43
LDA	33.92 (1.18)	33.95 (1.27)	0.02	20.54 (1.14)	20.55 (0.90)	0.08	15.51 (1.28)	15.65 (1.59)	0.62	28.74 (0.94)	28.63 (0.96)	0.02	17.54 (0.95)	17.40 (0.88)	0.10	11.53 (0.78)	11.33 (0.90)	0.94
1NN	45.02 (1.33)	45.05 (1.18)	0.01	40.65 (1.48)	40.35 (1.31)	0.01	35.70 (1.28)	35.60 (1.27)	0.05	41.33 (0.95)	41.40 (0.75)	0.01	35.59 (1.28)	35.77 (1.12)	0.02	32.57 (1.01)	32.57 (1.02)	0.11
nBayes	33.98 (1.20)	34.05 (0.88)	0.12	18.73 (0.91)	18.65 (0.80)	0.42	7.67 (0.78)	7.65 (0.72)	2.01	28.33 (0.77)	28.37 (0.72)	0.16	15.91 (0.82)	15.90 (0.80)	0.75	7.64 (0.50)	7.63 (0.46)	3.03
GBDT	8.20 (0.65)	8.20 (0.62)	5.25	1.76 (0.47)	1.80 (0.52)	8.27	0.32 (0.17)	0.30 (0.17)	35.79	9.30 (0.68)	9.37 (0.51)	5.67	3.06 (0.59)	3.03 (0.55)	20.18	0.84 (0.28)	0.80 (0.27)	50.56
AdaBoost	9.11 (0.89)	9.15 (0.84)	6.13	2.29 (0.56)	2.10 (0.41)	11.77	0.64 (0.26)	0.60 (0.32)	51.55	8.35 (0.61)	8.27 (0.61)	6.21	2.25 (0.31)	2.27 (0.35)	15.20	0.94 (0.25)	0.93 (0.25)	60.13
XGBoost	40.98 (1.92)	41.15 (2.29)	0.02	40.91 (1.84)	40.60 (2.41)	0.02	39.47 (1.93)	38.85 (1.47)	0.03	30.37 (1.05)	30.33 (0.97)	0.01	30.25 (1.19)	30.13 (1.42)	0.02	30.49 (0.90)	30.43 (0.66)	0.04

Table S24: Misclassification rates multiplied by 100 for Example 3 with balanced and imbalanced settings

	n = 50 (balanced)									n = 50 (imbalanced)								
	p = 50			p = 200			p = 1000			p = 50			p = 200			p = 1000		
	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time	Mean(SD)	Median(RSD)	Time
Multimodal(ROT)	0.33 (0.47)	0.00 (0.75)	1.48	0.00 (0.00)	0.00 (0.00)	2.95	0.00 (0.00)	0.00 (0.00)	10.88	0.68 (0.57)	0.67 (1.00)	1.75	0.00 (0.00)	0.00 (0.00)	3.41	0.00 (0.00)	0.00 (0.00)	12.47
Multimodal(NS)	0.45 (0.60)	0.00 (0.75)	1.30	0.00 (0.00)	0.00 (0.00)	2.25	0.00 (0.00)	0.00 (0.00)	9.19	0.70 (0.56)	0.67 (1.00)	1.44	0.02 (0.11)	0.00 (0.00)	2.60	0.00 (0.00)	0.00 (0.00)	11.27
Centroid	37.78 (4.58)	38.00 (3.73)	0.01	41.23 (4.31)	41.00 (3.73)	0.01	35.67 (4.62)	37.00 (5.60)	0.01	38.15 (3.03)	38.00 (2.74)	0.01	35.03 (3.08)	35.33 (2.99)	0.01	31.67 (3.15)	32.33 (3.98)	0.01
Median	36.20 (4.80)	35.50 (4.48)	0.01	33.00 (4.30)	34.00 (4.48)	0.04	25.25 (4.52)	26.00 (2.80)	0.20	35.27 (3.46)	35.67 (2.36)	0.01	26.37 (3.10)	26.33 (3.23)	0.06	20.30 (2.92)	20.00 (2.24)	0.19
Quantile	3.98 (3.32)	3.50 (3.73)	0.62	3.45 (2.65)	3.00 (2.24)	2.18	1.69 (1.44)	1.00 (1.49)	9.91	1.97 (1.77)	1.33 (1.40)	0.65	1.58 (1.55)	1.33 (0.10)	2.14	0.35 (0.52)	0.00 (0.50)	9.94
SVM	8.55 (2.29)	9.00 (2.99)	0.01	2.03 (1.05)	2.00 (1.49)	0.02	0.00 (0.00)	0.00 (0.00)	0.12	9.72 (2.05)	9.67 (2.11)	0.01	2.72 (1.24)	2.67 (1.00)	0.02	0.00 (0.00)	0.00 (0.00)	0.14
EQC	2.00 (0.00)	2.00 (0.00)	0.20	0.00 (0.00)	0.00 (0.00)	0.67	0.00 (0.00)	0.00 (0.00)	3.01	0.00 (0.00)	0.00 (0.00)	0.21	0.65 (0.11)	0.67 (0.00)	0.72	0.00 (0.00)	0.00 (0.00)	3.12
LDA	40.83 (7.53)	40.50 (7.09)	0.01	20.03 (5.22)	19.50 (3.73)	0.03	11.69 (3.33)	12.00 (2.99)	0.30	19.48 (3.63)	19.33 (2.99)	0.01	21.87 (3.68)	21.33 (4.10)	0.04	10.57 (2.87)	10.00 (2.61)	0.31
1-NN	40.03 (3.90)	40.00 (3.92)	0.01	46.38 (3.50)	46.00 (3.17)	0.01	42.23 (4.48)	42.00 (2.99)	0.01	44.92 (3.35)	46.00 (3.11)	0.01	37.85 (3.20)	38.00 (3.61)	0.01	43.15 (3.77)	43.33 (3.11)	0.01
nBayes	28.90 (4.34)	29.00 (3.73)	0.04	27.88 (3.37)	27.50 (4.48)	0.13	20.08 (3.30)	19.00 (1.49)	0.61	22.30 (3.07)	22.33 (3.61)	0.06	24.13 (2.65)	24.33 (2.61)	0.17	17.85 (2.37)	18.00 (2.49)	0.90
AdaBoost	0.90 (0.90)	1.00 (0.75)	4.07	0.40 (0.63)	0.00 (0.75)	7.74	0.17 (0.39)	0.00 (0.00)	32.49	0.53 (0.57)	0.67 (0.50)	4.22	0.30 (0.45)	0.00 (0.50)	8.03	0.23 (0.36)	0.00 (0.50)	33.54
GBDT	0.03 (0.16)	0.00 (0.00)	3.63	0.00 (0.00)	0.00 (0.00)	4.72	0.00 (0.00)	0.00 (0.00)	17.89	0.00 (0.00)	0.00 (0.00)	4.02	0.00 (0.00)	0.00 (0.00)	4.20	0.00 (0.00)	0.00 (0.00)	17.86
XGBoost	10.15 (4.21)	10.00 (3.92)	0.01	11.38 (4.10)	11.00 (3.73)	0.02	14.00 (2.49)	14.50 (2.05)	0.02	6.70 (2.81)	6.67 (3.11)	0.01	7.72 (2.04)	7.67 (2.49)	0.02	8.20 (1.92)	8.00 (1.99)	0.02
	n = 100 (balanced)									n = 100 (imbalanced)								
Multimodal(ROT)	0.10 (0.14)	0.00 (0.19)	1.85	0.00 (0.00)	0.00 (0.00)	3.45	0.00 (0.00)	0.00 (0.00)	11.60	0.56 (0.28)	0.50 (0.25)	2.45	0.00 (0.00)	0.00 (0.00)	2.57	0.00 (0.00)	0.00 (0.00)	11.69
Unimodal(NS)	0.13 (0.17)	0.00 (0.19)	1.77	0.00 (0.00)	0.00 (0.00)	3.05	0.00 (0.00)	0.00 (0.00)	10.42	0.60 (0.28)	0.50 (0.28)	1.86	0.00 (0.00)	0.00 (0.00)	2.10	0.00 (0.00)	0.00 (0.00)	10.89
Centroid	34.56 (2.53)	35.00 (2.33)	0.01	32.46 (2.69)	32.38 (2.29)	0.01	24.59 (2.09)	24.38 (2.29)	0.01	36.10 (2.11)	36.33 (2.33)	0.01	30.25 (2.11)	30.08 (2.02)	0.01	19.55 (1.77)	19.67 (1.80)	0.01
Median	32.68 (2.85)	33.13 (2.85)	0.02	29.16 (2.48)	29.00 (2.99)	0.05	19.04 (2.59)	18.88 (2.29)	0.20	33.25 (3.12)	33.42 (2.49)	0.02	27.92 (2.48)	28.25 (2.52)	0.06	15.37 (1.69)	15.58 (1.55)	0.19
Quantile	0.51 (0.55)	0.25 (0.56)	0.69	0.20 (0.38)	0.00 (0.19)	2.13	0.29 (0.33)	0.25 (0.37)	9.82	0.49 (0.33)	0.50 (0.37)	0.74	0.12 (0.17)	0.00 (0.12)	2.38	0.14 (0.18)	0.08 (0.12)	10.30
EQC	1.49 (0.08)	1.50 (0.00)	0.23	0.00 (0.00)	0.00 (0.00)	0.73	0.00 (0.00)	0.00 (0.00)	3.13	0.00 (0.00)	0.00 (0.00)	0.24	0.00 (0.00)	0.00 (0.00)	0.73	0.00 (0.00)	0.00 (0.00)	3.54
SVM	7.28 (1.10)	7.25 (0.98)	0.01	1.83 (0.65)	1.88 (0.61)	0.02	0.00 (0.00)	0.00 (0.00)	0.12	7.06 (0.91)	7.08 (0.93)	0.01	1.75 (0.51)	1.67 (0.50)	0.03	0.00 (0.00)	0.00 (0.00)	0.19
LDA	10.83 (1.28)	10.75 (1.49)	0.01	8.89 (1.21)	8.75 (0.98)	0.04	9.89 (1.95)	9.75 (1.87)	0.37	8.16 (0.85)	8.33 (0.44)	0.02	4.12 (0.74)	4.00 (0.65)	0.05	8.94 (1.77)	8.92 (1.68)	0.47
1-NN	45.58 (1.97)	45.63 (1.59)	0.01	46.63 (1.96)	46.25 (2.29)	0.01	46.99 (1.85)	47.13 (1.91)	0.01	41.45 (1.43)	41.25 (1.21)	0.01	40.83 (1.77)	40.83 (2.02)	0.01	38.76 (1.44)	38.92 (1.68)	0.03
nBayes	19.32 (2.63)	19.63 (3.26)	0.07	15.01 (1.67)	15.13 (1.73)	0.24	11.06 (1.94)	10.50 (1.87)	1.11	16.46 (2.02)	16.83 (1.80)	0.11	11.99 (1.39)	12.25 (1.43)	0.32	6.24 (0.98)	6.17 (0.87)	1.63
AdaBoost	0.00 (0.00)	0.00 (0.00)	4.20	0.00 (0.00)	0.00 (0.00)	9.01	0.00 (0.00)	0.00 (0.00)	34.35	0.00 (0.00)	0.00 (0.00)	4.70	0.00 (0.00)	0.00 (0.00)	9.25	0.00 (0.00)	0.00 (0.00)	39.04
GBDT	0.00 (0.00)	0.00 (0.00)	4.43	0.00 (0.00)	0.00 (0.00)	4.74	0.00 (0.00)	0.00 (0.00)	21.99	0.00 (0.00)	0.00 (0.00)	5.02	0.00 (0.00)	0.00 (0.00)	6.25	0.00 (0.00)	0.00 (0.00)	28.29
XGBoost	3.01 (1.49)	2.50 (1.68)	0.01	3.48 (1.56)	3.00 (1.68)	0.02	4.05 (1.66)	3.88 (1.73)	0.02	2.27 (0.92)	2.17 (0.90)	0.01	2.78 (1.06)	2.50 (0.65)	0.02	3.57 (0.71)	3.67 (0.62)	0.03
	n = 200 (balanced)									n = 200 (imbalanced)								
Multimodal(ROT)	0.19 (0.10)	0.20 (0.09)	1.95	0.00 (0.00)	0.00 (0.00)	3.73	0.00 (0.00)	0.00 (0.00)	11.31	0.33 (0.11)	0.33 (0.11)	0.97	0.00 (0.00)	0.00 (0.00)	2.87	0.00 (0.00)	0.00 (0.00)	12.95
Multimodal(NS)	0.21 (0.10)	0.20 (0.07)	1.77	0.00 (0.00)	0.00 (0.00)	3.37	0.00 (0.00)	0.00 (0.00)	10.85	0.37 (0.12)	0.40 (0.15)	1.32	0.00 (0.00)	0.00 (0.00)	2.57	0.00 (0.00)	0.00 (0.00)	11.50
Centroid	34.11 (1.96)	34.00 (1.53)	0.01	28.28 (1.85)	28.55 (1.74)	0.01	17.78 (1.50)	17.45 (1.36)	0.01	33.71 (1.56)	33.60 (1.39)	0.01	26.81 (1.31)	26.67 (1.52)	0.01	15.76 (0.92)	15.60 (0.68)	0.01
Median	32.61 (2.16)	32.95 (2.48)	0.02	26.74 (2.11)	26.55 (2.87)	0.05	15.61 (1.61)	15.50 (1.21)	0.21	31.47 (1.81)	31.77 (1.82)	0.02	23.75 (2.00)	23.63 (1.43)	0.05	14.50 (1.42)	14.70 (1.60)	0.22
Quantile	0.26 (0.14)	0.25 (0.09)	0.91	0.09 (0.12)	0.10 (0.07)	2.51	0.05 (0.09)	0.00 (0.07)	10.58	0.30 (0.15)	0.27 (0.15)	0.91	0.05 (0.06)	0.00 (0.05)	2.61	0.05 (0.07)	0.00 (0.05)	11.15
EQC	0.10 (0.02)	0.10 (0.00)	0.24	0.00 (0.00)	0.00 (0.00)	0.78	0.00 (0.00)	0.00 (0.00)	3.62	0.48 (0.07)	0.47 (0.00)	0.29	0.00 (0.00)	0.00 (0.00)	0.90	0.00 (0.00)	0.00 (0.00)	4.65
SVM	7.69 (0.71)	7.85 (0.69)	0.02	1.52 (0.26)	1.55 (0.24)	0.04	0.00 (0.00)	0.00 (0.00)	0.22	5.87 (0.39)	5.93 (0.41)	0.02	1.49 (0.23)	1.47 (0.30)	0.06	0.03 (0.01)	0.00 (0.00)	0.40
LDA	8.87 (0.77)	8.80 (0.58)	0.01	3.28 (0.51)	3.30 (0.47)	0.09	1.08 (0.33)	1.05 (0.30)	0.63	6.59 (0.43)	6.50 (0.36)	0.02	2.32 (0.32)	2.27 (0.31)	0.09	0.40 (0.14)	0.40 (0.15)	0.97
1-NN	47.62 (1.21)	47.65 (1.06)	0.01	45.80 (1.45)	45.80 (2.00)	0.01	43.55 (1.45)	43.55 (1.60)	0.05	39.98 (0.80)	40.07 (0.91)	0.01	40.00 (1.09)	40.20 (1.12)	0.02	38.49 (1.14)	38.47 (1.27)	0.13
nBayes	14.59 (1.49)	14.55 (1.23)	0.12	10.07 (1.16)	10.15 (1.29)	0.67	5.85 (0.80)	5.90 (0.90)	2.13	12.82 (1.10)	12.77 (1.07)	0.15	7.53 (1.09)	7.43 (1.16)	0.58	4.13 (0.63)	4.10 (0.58)	3.03
AdaBoost	0.00 (0.00)	0.00 (0.00)	4.70	0.00 (0.00)	0.00 (0.00)	9.70	0.00 (0.00)	0.00 (0.00)	41.42	0.01 (0.03)	0.00 (0.00)	5.26	0.00 (0.00)	0.00 (0.00)	11.65	0.00 (0.00)	0.00 (0.00)	53.93
GBDT	0.00 (0.00)	0.00 (0.00)	5.43	0.00 (0.00)	0.00 (0.00)	8.11	0.00 (0.00)	0.00 (0.00)	37.71	0.00 (0.00)	0.00 (0.00)	5.31	0.00 (0.00)	0.00 (0.00)	20.07	0.00 (0.00)	0.00 (0.00)	50.74
XGBoost	2.92 (0.53)	2.95 (0.52)	0.01	2.90 (0.58)	2.80 (0.56)	0.02	2.82 (0.66)	2.80 (0.47)	0.03	1.84 (0.32)	1.80 (0.31)	0.01	2.01 (0.37)	1.97 (0.31)	0.02	2.01 (0.36)	2.00 (0.31)	0.04

Table S25: Descriptive statistics for the number of modes for $p = 6033$ genes determined by the three testing procedures in the prostate cancer dataset, nominal level $\alpha = 0.05$

Test	Group	Min	Q_1	Meidan	Mean	Q_3	Max
SI	healthy	1.000	1.000	1.000	1.000	1.000	1.000
	cancer	1.000	1.000	1.000	1.001	1.000	2.000
FM	healthy	1.000	1.000	17.00	17.47	25.00	123.00
	cancer	1.000	1.000	18.00	18.83	26.00	150.00
ACR	healthy	1.000	1.000	1.000	1.516	2.000	10.000
	cancer	1.000	1.000	1.000	1.543	2.000	7.000

Table S26: Testing results for the number of modes for $p = 50, 200, 500, 1000$ selected predictors of two groups based on three testing procedures for the second dataset, critical value $\alpha = 0.05$

		no-focal-lesion group				with-focal-lesion group			
p	Test	Mean	Median	Min	Max	Mean	Median	Min	Max
$p = 50$	SI	1.000	1.00	1.00	1.00	1.020	1.00	1.00	2.00
	FM	1.250	1.00	1.00	3.00	2.700	1.00	1.00	58.00
	ACR	1.065	1.00	1.00	2.00	1.075	1.00	1.00	3.00
$p = 200$	SI	1.000	1.00	1.00	1.00	1.000	1.00	1.00	1.00
	FM	1.195	1.00	1.00	9.00	1.995	1.00	1.00	58.00
	ACR	1.160	1.00	1.00	2.00	1.220	1.00	1.00	3.00
$p = 500$	SI	1.000	1.00	1.00	1.00	1.000	1.00	1.00	1.00
	FM	1.480	1.00	1.00	119.00	2.150	1.00	1.00	106.00
	ACR	1.064	1.00	1.00	4.00	1.068	1.00	1.00	3.00
$p = 1000$	SI	1.000	1.00	1.00	1.00	1.000	1.00	1.00	1.00
	FM	1.696	1.00	1.00	119.00	2.216	1.00	1.00	210.00
	ACR	1.063	1.00	1.00	4.00	1.080	1.00	1.00	4.00

Table S27: Mean misclassification error rates (%) for the multiple myeloma dataset dataset, with standard errors (%) in parenthese

Classifiers	$p = 50$	$p = 200$	$p = 500$	$p = 1000$
Unimodal classifier (no test)	10.50 (2.55)	7.03 (3.76)	3.46 (2.98)	1.70 (2.74)
Bimodal classifier (no test)	40.54 (13.44)	31.43 (8.65)	35.40 (13.30)	34.90 (8.31)
Mode-based classifier (SI test)	18.68 (5.78)	7.03 (3.76)	3.46 (2.98)	1.70 (2.74)
Mode-based classifier (FM test)	21.38 (9.34)	18.02 (10.17)	13.93 (6.81)	9.88 (7.34)
Mode-based classifier (ACR test)	19.13 (7.69)	15.66 (9.13)	14.66 (8.65)	13.00 (10.16)
Quantile-mode classifier	21.50 (7.08)	20.79 (7.90)	14.62 (8.52)	10.50 (8.95)
Centroid classifier	34.87 (13.44)	36.14 (14.43)	34.87 (8.19)	37.60 (9.83)
Median classifier	30.92 (14.59)	31.50 (16.96)	30.23 (9.06)	27.63 (13.01)
Quantile classifier	23.89 (10.18)	26.27 (6.66)	23.24 (9.83)	24.95 (5.94)
Ensemble quantile classifier	17.94 (10.94)	20.86 (7.01)	16.18 (7.43)	17.32 (8.44)
Naive Bayes classifier	26.76 (14.15)	26.21 (10.52)	23.33 (14.5)	26.62 (10.58)
LDA	25.00 (11.08)	32.03(9.40)	16.27 (9.05)	19.93 (12.93)
1-NN	25.03 (11.84)	29.74 (10.33)	34.31 (7.53)	32.48 (10.69)
SVM	26.14 (13.02)	29.77 (13.54)	23.10 (9.31)	33.12 (5.02)
AdaBoost	20.92 (10.02)	21.63 (10.12)	19.12 (8.07)	19.69 (12.51)
GBDT	19.71 (12.23)	19.90 (14.48)	22.22 (19.43)	22.66 (22.52)
XGBoost	23.82 (8.79)	29.15 (8.54)	23.86 (7.09)	21.38 (6.17)